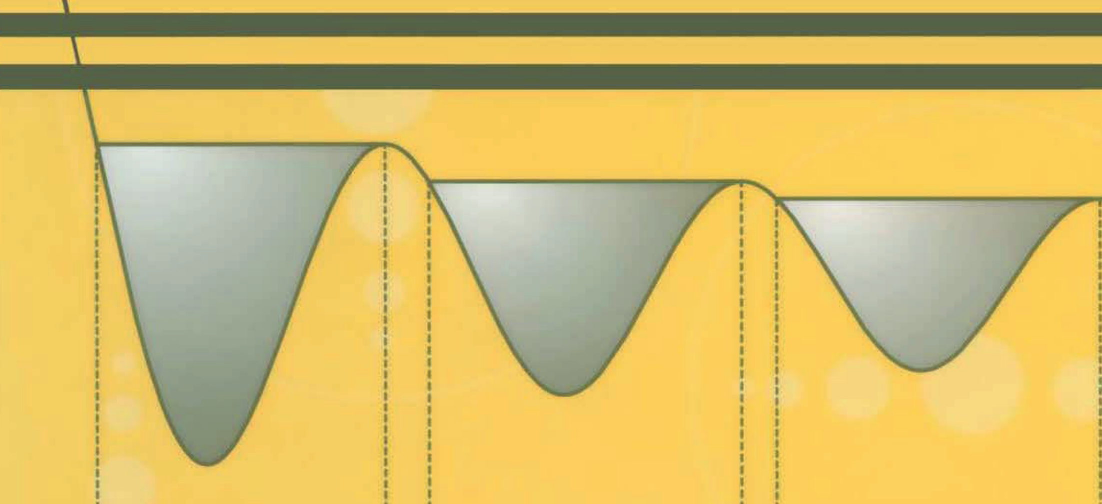
CORNERSTONES

BASIC REAL ANALYSIS

Anthony W. Knapp



Birkhäuser

Cornerstones

Anthony W. Knapp

Basic Real Analysis

Along with a companion volume
Advanced Real Analysis

Birkhäuser
Boston • Basel • Berlin

Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, NY 11733-1729
U.S.A.
e-mail to: aknapp@math.sunysb.edu
http://www.math.sunysb.edu/~aknapp/books/basic.html

Cover design by Mary Burgess.

Mathematics Subject Classicification (2000): 28-01, 26-01, 42-01, 54-01, 34-01

Library of Congress Cataloging-in-Publication Data
Knapp, Anthony W.
Basic real analysis: along with a companion volume Advanced real analysis / Anthony W. Knapp
p. cm. – (Cornerstones)
Includes bibliographical references and index.
ISBN 0-8176-3250-6 (alk. paper)
1. Mathematical analysis. I. Title. II. Cornerstones (Birkhäuser)

QA300.K56 2005
515–dc22 2005048070

ISBN-10 0-8176-3250-6 eISBN 0-8176-4441-5 Printed on acid-free paper.
ISBN-13 978-0-8176-3250-2

Advanced Real Analysis ISBN 0-8176-4382-6
Basic Real Analysis and *Advanced Real Analysis* (Set) ISBN 0-8176-4407-5

Printed in the United States of America. (MP)

9 8 7 6 5 4 3 2 1 SPIN 10934074

www.birkhauser.com

To Susan

and

To My Real-Analysis Teachers:

Salomon Bochner, William Feller, Hillel Furstenberg,
Harish-Chandra, Sigurdur Helgason, John Kemeny,
John Lamperti, Hazleton Mirkil, Edward Nelson,
Laurie Snell, Elias Stein, Richard Williamson

CONTENTS

Preface xi
Dependence Among Chapters xiv
Guide for the Reader xv
List of Figures xviii
Acknowledgments xix
Standard Notation xxi

I. THEORY OF CALCULUS IN ONE REAL VARIABLE 1

1. Review of Real Numbers, Sequences, Continuity 2
2. Interchange of Limits 13
3. Uniform Convergence 15
4. Riemann Integral 26
5. Complex-Valued Functions 41
6. Taylor's Theorem with Integral Remainder 43
7. Power Series and Special Functions 44
8. Summability 53
9. Weierstrass Approximation Theorem 58
10. Fourier Series 61
11. Problems 78

II. METRIC SPACES 82

1. Definition and Examples 83
2. Open Sets and Closed Sets 91
3. Continuous Functions 95
4. Sequences and Convergence 97
5. Subspaces and Products 102
6. Properties of Metric Spaces 105
7. Compactness and Completeness 108
8. Connectedness 115
9. Baire Category Theorem 117
10. Properties of $C(S)$ for Compact Metric S 121
11. Completion 127
12. Problems 130

III. THEORY OF CALCULUS IN SEVERAL REAL VARIABLES 135
1. Operator Norm 135
2. Nonlinear Functions and Differentiation 139
3. Vector-Valued Partial Derivatives and Riemann Integrals 146
4. Exponential of a Matrix 148
5. Partitions of Unity 151
6. Inverse and Implicit Function Theorems 152
7. Definition and Properties of Riemann Integral 161
8. Riemann Integrable Functions 166
9. Fubini's Theorem for the Riemann Integral 169
10. Change of Variables for the Riemann Integral 171
11. Problems 179

IV. THEORY OF ORDINARY DIFFERENTIAL EQUATIONS AND SYSTEMS 183
1. Qualitative Features and Examples 183
2. Existence and Uniqueness 187
3. Dependence on Initial Conditions and Parameters 194
4. Integral Curves 199
5. Linear Equations and Systems, Wronskian 201
6. Homogeneous Equations with Constant Coefficients 208
7. Homogeneous Systems with Constant Coefficients 211
8. Series Solutions in the Second-Order Linear Case 218
9. Problems 226

V. LEBESGUE MEASURE AND ABSTRACT MEASURE THEORY 231
1. Measures and Examples 231
2. Measurable Functions 238
3. Lebesgue Integral 241
4. Properties of the Integral 245
5. Proof of the Extension Theorem 253
6. Completion of a Measure Space 262
7. Fubini's Theorem for the Lebesgue Integral 265
8. Integration of Complex-Valued and Vector-Valued Functions 274
9. L^1, L^2, L^∞, and Normed Linear Spaces 279
10. Problems 289

VI. MEASURE THEORY FOR EUCLIDEAN SPACE 296
1. Lebesgue Measure and Other Borel Measures 297
2. Convolution 306
3. Borel Measures on Open Sets 314
4. Comparison of Riemann and Lebesgue Integrals 318

VI. MEASURE THEORY FOR EUCLIDEAN SPACE (Continued)
5. Change of Variables for the Lebesgue Integral 320
6. Hardy–Littlewood Maximal Theorem 327
7. Fourier Series and the Riesz–Fischer Theorem 334
8. Stieltjes Measures on the Line 339
9. Fourier Series and the Dirichlet–Jordan Theorem 346
10. Distribution Functions 350
11. Problems 352

VII. DIFFERENTIATION OF LEBESGUE INTEGRALS ON THE LINE 357
1. Differentiation of Monotone Functions 357
2. Absolute Continuity, Singular Measures, and Lebesgue Decomposition 364
3. Problems 370

VIII. FOURIER TRANSFORM IN EUCLIDEAN SPACE 373
1. Elementary Properties 373
2. Fourier Transform on L^1, Inversion Formula 377
3. Fourier Transform on L^2, Plancherel Formula 381
4. Schwartz Space 384
5. Poisson Summation Formula 389
6. Poisson Integral Formula 392
7. Hilbert Transform 397
8. Problems 404

IX. L^p SPACES 409
1. Inequalities and Completeness 409
2. Convolution Involving L^p 417
3. Jordan and Hahn Decompositions 418
4. Radon–Nikodym Theorem 420
5. Continuous Linear Functionals on L^p 424
6. Marcinkiewicz Interpolation Theorem 427
7. Problems 436

X. TOPOLOGICAL SPACES 441
1. Open Sets and Constructions of Topologies 441
2. Properties of Topological Spaces 447
3. Compactness and Local Compactness 451
4. Product Spaces and the Tychonoff Product Theorem 458
5. Sequences and Nets 463
6. Quotient Spaces 471

X. TOPOLOGICAL SPACES (Continued)

7. Urysohn's Lemma 474
8. Metrization in the Separable Case 476
9. Ascoli–Arzelà and Stone–Weierstrass Theorems 477
10. Problems 480

XI. INTEGRATION ON LOCALLY COMPACT SPACES 485

1. Setting 485
2. Riesz Representation Theorem 490
3. Regular Borel Measures 504
4. Dual to Space of Finite Signed Measures 509
5. Problems 517

XII. HILBERT AND BANACH SPACES 520

1. Definitions and Examples 520
2. Geometry of Hilbert Space 526
3. Bounded Linear Operators on Hilbert Spaces 535
4. Hahn–Banach Theorem 537
5. Uniform Boundedness Theorem 543
6. Interior Mapping Principle 545
7. Problems 549

APPENDIX 553

A1. Sets and Functions 553
A2. Mean Value Theorem and Some Consequences 559
A3. Inverse Function Theorem in One Variable 561
A4. Complex Numbers 563
A5. Classical Schwarz Inequality 563
A6. Equivalence Relations 564
A7. Linear Transformations, Matrices, and Determinants 565
A8. Factorization and Roots of Polynomials 568
A9. Partial Orderings and Zorn's Lemma 573
A10. Cardinality 577

Hints for Solutions of Problems 581
Selected References 637
Index of Notation 639
Index 643

PREFACE

This book and its companion volume *Advanced Real Analysis* systematically develop concepts and tools in real analysis that are vital to every mathematician, whether pure or applied, aspiring or established. The two books together contain what the young mathematician needs to know about real analysis in order to communicate well with colleagues in all branches of mathematics.

The books are written as textbooks, and their primary audience is students who are learning the material for the first time and who are planning a career in which they will use advanced mathematics professionally. Much of the material in the books corresponds to normal course work. Nevertheless, it is often the case that core mathematics curricula, time-limited as they are, do not include all the topics that one might like. Thus the book includes important topics that may be skipped in required courses but that the professional mathematician will ultimately want to learn by self-study.

The content of the required courses at each university reflects expectations of what students need before beginning specialized study and work on a thesis. These expectations vary from country to country and from university to university. Even so, there seems to be a rough consensus about what mathematics a plenary lecturer at a broad international or national meeting may take as known by the audience. The tables of contents of the two books represent my own understanding of what that degree of knowledge is for real analysis today.

Key topics and features of *Basic Real Analysis* are as follows:

- Early chapters treat the fundamentals of real variables, sequences and series of functions, the theory of Fourier series for the Riemann integral, metric spaces, and the theoretical underpinnings of multivariable calculus and ordinary differential equations.
- Subsequent chapters develop the Lebesgue theory in Euclidean and abstract spaces, Fourier series and the Fourier transform for the Lebesgue integral, point-set topology, measure theory in locally compact Hausdorff spaces, and the basics of Hilbert and Banach spaces.
- The subjects of Fourier series and harmonic functions are used as recurring motivation for a number of theoretical developments.
- The development proceeds from the particular to the general, often introducing examples well before a theory that incorporates them.

- More than 300 problems at the ends of chapters illuminate aspects of the text, develop related topics, and point to additional applications. A separate 55-page section "Hints for Solutions of Problems" at the end of the book gives detailed hints for most of the problems, together with complete solutions for many.

Beyond a standard calculus sequence in one and several variables, the most important prerequisite for using *Basic Real Analysis* is that the reader already know what a proof is, how to read a proof, and how to write a proof. This knowledge typically is obtained from honors calculus courses, or from a course in linear algebra, or from a first junior-senior course in real variables. In addition, it is assumed that the reader is comfortable with a modest amount of linear algebra, including row reduction of matrices, vector spaces and bases, and the associated geometry. A passing acquaintance with the notions of group, subgroup, and quotient is helpful as well.

Chapters I–IV are appropriate for a single rigorous real-variables course and may be used in either of two ways. For students who have learned about proofs from honors calculus or linear algebra, these chapters offer a full treatment of real variables, leaving out only the more familiar parts near the beginning—such as elementary manipulations with limits, convergence tests for infinite series with positive scalar terms, and routine facts about continuity and differentiability. For students who have learned about proofs from a first junior-senior course in real variables, these chapters are appropriate for a second such course that begins with Riemann integration and sequences and series of functions; in this case the first section of Chapter I will be a review of some of the more difficult foundational theorems, and the course can conclude with an introduction to the Lebesgue integral from Chapter V if time permits.

Chapters V through XII treat Lebesgue integration in various settings, as well as introductions to the Euclidean Fourier transform and to functional analysis. Typically this material is taught at the graduate level in the United States, frequently in one of three ways: The first way does Lebesgue integration in Euclidean and abstract settings and goes on to consider the Euclidean Fourier transform in some detail; this corresponds to Chapters V–VIII. A second way does Lebesgue integration in Euclidean and abstract settings, treats L^p spaces and integration on locally compact Hausdorff spaces, and concludes with an introduction to Hilbert and Banach spaces; this corresponds to Chapters V–VII, part of IX, and XI–XII. A third way combines an introduction to the Lebesgue integral and the Euclidean Fourier transform with some of the subject of partial differential equations; this corresponds to some portion of Chapters V–VI and VIII, followed by chapters from the companion volume *Advanced Real Analysis*.

In my own teaching, I have most often built one course around Chapters I–IV and another around Chapters V–VII, part of IX, and XI–XII. I have normally assigned the easier sections of Chapters II and X as outside reading, indicating the date when the lectures would begin to use that material.

More detailed information about how the book may be used with courses may be deduced from the chart "Dependence among Chapters" on page xiv and the section "Guide to the Reader" on pages xv–xvii.

The problems at the ends of chapters are an important part of the book. Some of them are really theorems, some are examples showing the degree to which hypotheses can be stretched, and a few are just exercises. The reader gets no indication which problems are of which type, nor of which ones are relatively easy. Each problem can be solved with tools developed up to that point in the book, plus any additional prerequisites that are noted.

Two omissions from the pair of books are of note. One is any treatment of Stokes's Theorem and differential forms. Although there is some advantage, when studying these topics, in having the Lebesgue integral available and in having developed an attitude that integration can be defined by means of suitable linear functionals, the topic of Stokes's Theorem seems to fit better in a book about geometry and topology, rather than in a book about real analysis.

The other omission concerns the use of complex analysis. It is tempting to try to combine real analysis and complex analysis into a single subject, but my own experience is that this combination does not work well at the level of *Basic Real Analysis*, only at the level of *Advanced Real Analysis*.

Almost all of the mathematics in the two books is at least forty years old, and I make no claim that any result is new. The books are a distillation of lecture notes from a 35-year period of my own learning and teaching. Sometimes a problem at the end of a chapter or an approach to the exposition may not be a standard one, but no attempt has been made to identify such problems and approaches. In the reverse direction it is possible that my early lecture notes have directly quoted some source without proper attribution. As an attempt to rectify any difficulties of this kind, I have included a section of "Acknowledgements" on pages xix–xx of this volume to identify the main sources, as far as I can reconstruct them, for those original lecture notes.

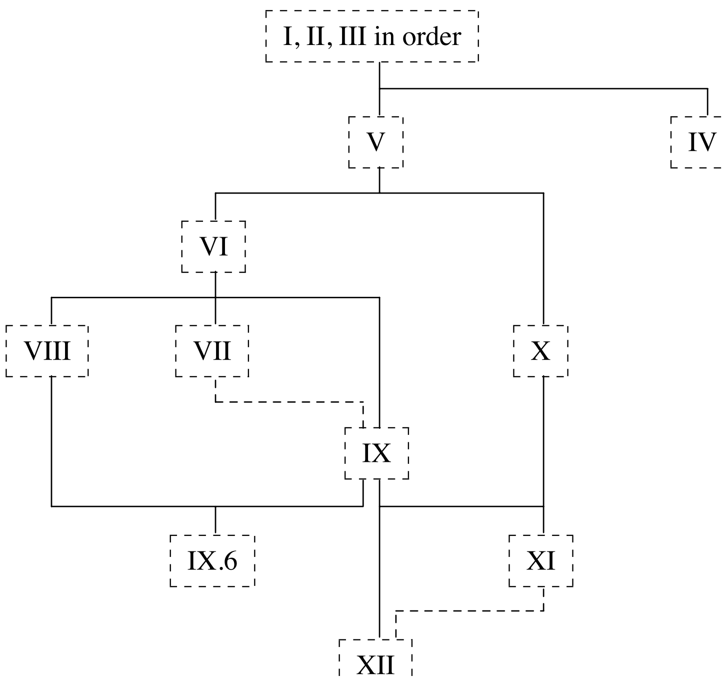
I am grateful to Ann Kostant and Steven Krantz for encouraging this project and for making many suggestions about pursuing it, and to Susan Knapp and David Kramer for helping with the readability. The typesetting was by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TeX, and the figures were drawn with Mathematica.

I invite corrections and other comments from readers. I plan to maintain a list of known corrections on my own Web page.

A. W. KNAPP
May 2005

DEPENDENCE AMONG CHAPTERS

Below is a chart of the main lines of dependence of chapters on prior chapters. The dashed lines indicate helpful motivation but no logical dependence. Apart from that, particular examples may make use of information from earlier chapters that is not indicated by the chart.



GUIDE FOR THE READER

This section is intended to help the reader find out what parts of each chapter are most important and how the chapters are interrelated. Further information of this kind is contained in the abstracts that begin each of the chapters.

The book pays attention to certain recurring themes in real analysis, allowing a person to see how these themes arise in increasingly sophisticated ways. Examples are the role of interchanges of limits in theorems, the need for certain explicit formulas in the foundations of subject areas, the role of compactness and completeness in existence theorems, and the approach of handling nice functions first and then passing to general functions.

All of these themes are introduced in Chapter I, and already at that stage they interact in subtle ways. For example, a natural investigation of interchanges of limits in Sections 2–3 leads to the discovery of Ascoli's Theorem, which is a fundamental compactness tool for proving existence results. Ascoli's Theorem is proved by the "Cantor diagonal process," which has other applications to compactness questions and does not get fully explained until Chapter X. The consequence is that, no matter where in the book a reader plans to start, everyone will be helped by at least leafing through Chapter I.

The remainder of this section is an overview of individual chapters and groups of chapters.

Chapter I. Every section of this chapter plays a role in setting up matters for later chapters. No knowledge of metric spaces is assumed anywhere in the chapter. Section 1 will be a review for anyone who has already had a course in real-variable theory; the section shows how compactness and completeness address all the difficult theorems whose proofs are often skipped in calculus. Section 2 begins the development of real-variable theory at the point of sequences and series of functions. It contains interchange results that turn out to be special cases of the main theorems of Chapter V. Sections 8–9 introduce the approach of handling nice functions before general functions, and Section 10 introduces Fourier series, which provided a great deal of motivation historically for the development of real analysis and are used in this book in that same way. Fourier series are somewhat limited in the setting of Chapter I because one encounters no class of functions, other than infinitely differentiable ones, that corresponds exactly to some class of Fourier coefficients; as a result Fourier series, with Riemann integration in use, are not particularly useful for constructing new functions from old ones. This defect will be fixed with the aid of the Lebesgue integral in Chapter VI.

Chapter II. Now that continuity and convergence have been addressed on the line, this chapter establishes a framework for these questions in higher-dimensional Euclidean space and other settings. There is no point in ad hoc definitions for each setting, and metric spaces handle many such settings at once. Chapter X later will enlarge the framework from metric spaces to "topological spaces." Sections 1–6 of Chapter II are routine. Section 7, on compactness and completeness, is the core. The Baire Category Theorem in Section 9 is not used outside of Chapter II until Chapter XII, and it may therefore be skipped temporarily. Section 10 contains the Stone–Weierstrass Theorem, which is a fundamental approximation tool. Section 11 is used in some of the problems but is not otherwise used in the book.

Chapter III. This chapter does for the several-variable theory what Chapter I has done for the one-variable theory. The main results are the Inverse and Implicit Function Theorems in Section 6 and the change-of-variables formula for multiple integrals in Section 10. The change-of-variables formula has to be regarded as only a preliminary version, since what it directly accomplishes for the change to polar coordinates still needs supplementing; this difficulty will be repaired in Chapter VI with the aid of the Lebesgue integral. Section 4, on exponentials of matrices, may be skipped if linear systems of ordinary differential equations are going to be skipped in Chapter IV. Some of the problems at the end of the chapter introduce harmonic functions; harmonic functions will be combined with Fourier series in problems in later chapters to motivate and illustrate some of the development.

Chapter IV provides theoretical underpinnings for the material in a traditional undergraduate course in ordinary differential equations. Nothing later in the book is logically dependent on Chapter IV; however, Chapter XII includes a discussion of orthogonal systems of functions, and the examples of these that arise in Chapter IV are helpful as motivation. Some people shy away from differential equations and might wish to treat Chapter IV only lightly, or perhaps not at all. The subject is nevertheless of great importance, and Chapter IV is the beginning of it. A minimal treatment of Chapter IV might involve Sections 1–2 and Section 8, all of which visibly continue the themes begun in Chapter I.

Chapters V–VI treat the core of measure theory—including the basic convergence theorems for integrals, the development of Lebesgue measure in one and several variables, Fubini's Theorem, the metric spaces L^1 and L^2 and L^∞, and the use of maximal theorems for getting at differentiation of integrals and other theorems concerning almost-everywhere convergence. In Chapter V Lebesgue measure in one dimension is introduced right away, so that one immediately has the most important example at hand. The fundamental Extension Theorem for getting measures to be defined on σ-rings and σ-algebras is stated when needed but is proved only after the basic convergence theorems for integrals have been proved; the proof in Sections 5–6 may be skipped on first reading. Section 7, on Fubini's Theorem, is a powerful result about interchange of integrals. At the same time that it justifies interchange, it also constructs a "double integral"; consequently the section prepares the way for the construction in Chapter VI of n-dimensional Lebesgue measure from 1-dimensional Lebesgue measure. Section 10 introduces normed linear spaces along with the examples of L^1 and L^2 and L^∞, and it goes on to establish some properties of all normed linear spaces. Chapter VI fleshes out measure theory as it applies to Euclidean space in more than one dimension. Of special note is the Lebesgue-integration version in Section 5 of the change-of-variables formula for multiple integrals and the Riesz–Fischer Theorem in Section 7. The latter characterizes square-integrable periodic functions by their Fourier coefficients and makes the subject of Fourier series useful in constructing functions. Differentiation of integrals in approached in Section 6 of Chapter VI as a problem of estimating finiteness of a quantity, rather than its smallness; the device is the Hardy–Littlewood Maximal Theorem, and the approach becomes a routine way of approaching almost-everywhere convergence theorems. Sections 8–10 are of somewhat less importance and may be omitted if time is short; Section 10 is applied only in Section IX.6.

Chapters VII–IX are continuations of measure theory that are largely independent of each other. Chapter VII contains the traditional proof of the differentiation of integrals on the line via differentiation of monotone functions. No later chapter is logically dependent on Chapter VII; the material is included only because of its historical importance and its usefulness as motivation for the Radon–Nikodym Theorem in Chapter IX. Chapter VIII is an introduction to the Fourier transform in Euclidean space. Its core consists of the first four sections, and the rest may be considered as optional if Section IX.6 is to be omitted. Chapter IX concerns L^p spaces for $1 \le p \le \infty$; only Section 6 makes use of material from Chapter VIII.

Chapter X develops, at the latest possible time in the book, the necessary part of point-set topology that goes beyond metric spaces. Emphasis is on product and quotient spaces, and on Urysohn's Lemma concerning the construction of real-valued functions on normal spaces.

Chapter XI contains one more continuation of measure theory, namely special features of measures on locally compact Hausdorff spaces. It provides an example beyond L^p spaces in which one can usefully identify the dual of a particular normed linear space. These chapters depend on Chapter X and on the first five sections of Chapter IX but do not depend on Chapters VII–VIII.

Chapter XII is a brief introduction to functional analysis, particularly to Hilbert spaces, Banach spaces, and linear operators on them. The main topics are the geometry of Hilbert space and the three main theorems about Banach spaces.

LIST OF FIGURES

1.1. Approximate identity 59
1.2. Fourier series of sawtooth function 65
1.3. Dirichlet kernel 69
2.1. An open set centered at the origin in the hedgehog space 88
2.2. Open ball contained in an intersection of two open balls 92
4.1. Graphs of solutions of some first-order ordinary differential equations 185
4.2. Integral curve of a vector field 199
4.3. Graph of Bessel function $J_0(t)$ 225
6.1. Construction of a Cantor function F 343
7.1. Rising Sun Lemma 358

ACKNOWLEDGMENTS

The author acknowledges the sources below as the main ones he used in preparing the lectures from which this book evolved. Any residual unattributed direct quotations in the book are likely to be from these.

The descriptions below have been abbreviated. Full descriptions of the books and Stone article may be found in the section "Selected References" at the end of the book. The item "Feller's *Functional Analysis*" refers to lectures by William Feller at Princeton University for Fall 1962 and Spring 1963, and the item "Nelson's *Probability*" refers to lectures by Edward Nelson at Princeton University for Spring 1963.

This list is not to be confused with a list of recommended present-day reading for these topics; newer books deserve attention.

CHAPTER I. Rudin's *Principles of Mathematical Analysis*, Zygmund's *Trigonometric Series*.

CHAPTER II. Feller's *Functional Analysis*, Kelley's *General Topology*, Stone's "A generalized Weierstrass approximation theorem."

CHAPTER III. Rudin's *Principles of Mathematical Analysis*, Spivak's *Calculus on Manifolds*.

CHAPTER IV. Coddington–Levinson's *Theory of Ordinary Differential Equations*, Kaplan's *Ordinary Differential Equations*.

CHAPTER V. Halmos's *Measure Theory*, Rudin's *Principles of Mathematical Analysis*.

CHAPTER VI. Rudin's *Principles of Mathematical Analysis*, Rudin's *Real and Complex Analysis*, Saks's *Theory of the Integral*, Spivak's *Calculus on Manifolds*, Stein–Weiss's *Introduction to Fourier Analysis on Euclidean Spaces*.

CHAPTER VII. Riesz–Nagy's *Functional Analysis*, Zygmund's *Trigonometric Series*.

CHAPTER VIII. Stein's *Singular Integrals and Differentiability Properties of Functions*, Stein–Weiss's *Introduction to Fourier Analysis on Euclidean Spaces*.

CHAPTER IX. Dunford–Schwartz's *Linear Operators*, Feller's *Functional Analysis*, Halmos's *Measure Theory*, Saks's *Theory of the Integral*, Stein's *Singular Integrals and Differentiability Properties of Functions*.

CHAPTER X. Kelley's *General Topology*, Nelson's *Probability*.

CHAPTER XI. Feller's *Functional Analysis*, Halmos's *Measure Theory*, Nelson's *Probability*.

CHAPTER XII. Dunford–Schwartz's *Linear Operators*, Feller's *Functional Analysis*, Riesz–Nagy's *Functional Analysis*.

APPENDIX. For Sections 1, 9, 10: Dunford–Schwartz's *Linear Operators*, Hayden–Kennison's *Zermelo–Fraenkel Set Theory*, Kelley's *General Topology*.

STANDARD NOTATION

Item	**Meaning**
$\#S$ or $\lvert S\rvert$	number of elements in S
$\varnothing$	empty set
$\{x \in E \mid P\}$	the set of x in E such that P holds
E^c	complement of the set E
$E \cup F,\ E \cap F,\ E - F$	union, intersection, difference of sets
$\bigcup_\alpha E_\alpha,\ \bigcap_\alpha E_\alpha$	union, intersection of the sets E_α
$E \subseteq F,\ E \supseteq F$	E is contained in F, E contains F
$E \times F,\ \bigtimes_{s\in S} X_s$	products of sets
$(a_1, \dots, a_n),\ \{a_1, \dots, a_n\}$	ordered n-tuple, unordered n-tuple
$f : E \to F,\ x \mapsto f(x)$	function, effect of function
$f \circ g,\ f\big\vert_E$	composition of f following g, restriction to E
$f(\,\cdot\,, y)$	the function $x \mapsto f(x, y)$
$f(E),\ f^{-1}(E)$	direct and inverse image of a set
δ_{ij}	Kronecker delta: 1 if $i = j$, 0 if $i \neq j$
$\binom{n}{k}$	binomial coefficient
n positive, n negative	$n > 0, n < 0$
$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$	integers, rationals, reals, complex numbers
max (and similarly min)	maximum of finite subset of a totally ordered set
$\sum$ or $\prod$	sum or product, possibly with a limit operation
countable	finite or in one-one correspondence with $\mathbb{Z}$
$[x]$	greatest integer $\leq x$ if x is real
$\operatorname{Re} z, \operatorname{Im} z$	real and imaginary parts of complex z
$\bar{z}$	complex conjugate of z
$\lvert z\rvert$	absolute value of z
1	multiplicative identity
1 or I	identity matrix or operator
$\dim V$	dimension of vector space
$\mathbb{R}^n, \mathbb{C}^n$	spaces of column vectors
$\det A$	determinant of A
A^{tr}	transpose of A
$\operatorname{diag}(a_1, \dots, a_n)$	diagonal matrix
$\cong$	is isomorphic to, is equivalent to

Basic Real Analysis

CHAPTER I

Theory of Calculus in One Real Variable

Abstract. This chapter, beginning with Section 2, develops the topic of sequences and series of functions, especially of functions of one variable. An important part of the treatment is an introduction to the problem of interchange of limits, both theoretically and practically. This problem plays a role repeatedly in real analysis, but its visibility decreases as more and more results are developed for handling it in various situations. Fourier series are introduced in this chapter and are carried along throughout the book as a motivating example for a number of problems in real analysis.

Section 1 makes contact with the core of a first undergraduate course in real-variable theory. Some material from such a course is repeated here in order to establish notation and a point of view. Omitted material is summarized at the end of the section, and some of it is discussed in a little more detail in an appendix at the end of the book. The point of view being established is the use of defining properties of the real number system to prove the Bolzano–Weierstrass Theorem, followed by the use of that theorem to prove some of the difficult theorems that are usually assumed in a one-variable calculus course. The treatment makes use of the extended real-number system, in order to allow sup and inf to be defined for any nonempty set of reals and to allow lim sup and lim inf to be meaningful for any sequence.

Sections 2–3 introduce the problem of interchange of limits. They show how certain concrete problems can be viewed in this way, and they give a way of thinking about all such interchanges in a common framework. A positive result affirms such an interchange under suitable hypotheses of monotonicity. This is by way of introduction to the topic in Section 3 of uniform convergence and the role of uniform convergence in continuity and differentiation.

Section 4 gives a careful development of the Riemann integral for real-valued functions of one variable, establishing existence of Riemann integrals for bounded functions that are discontinuous at only finitely many points, basic properties of the integral, the Fundamental Theorem of Calculus for continuous integrands, the change-of-variables formula, and other results. Section 5 examines complex-valued functions, pointing out the extent to which the results for real-valued functions in the first four sections extend to complex-valued functions.

Section 6 is a short treatment of the version of Taylor's Theorem in which the remainder is given by an integral. Section 7 takes up power series and uses them to define the elementary transcendental functions and establish their properties. The power series expansion of $(1+x)^p$ for arbitrary complex p is studied carefully. Section 8 introduces Cesàro and Abel summability, which play a role in the subject of Fourier series. A converse theorem to Abel's theorem is used to exhibit the function $|x|$ as the uniform limit of polynomials on $[-1, 1]$. The Weierstrass Approximation Theorem of Section 9 generalizes this example and establishes that every continuous complex-valued function on a closed bounded interval is the uniform limit of polynomials.

Section 10 introduces Fourier series in one variable in the context of the Riemann integral. The main theorems of the section are a convergence result for continuously differentiable functions, Bessel's inequality, the Riemann–Lebesgue Lemma, Fejér's Theorem, and Parseval's Theorem.

1. Review of Real Numbers, Sequences, Continuity

This section reviews some material that is normally in an undergraduate course in real analysis. The emphasis will be on a rigorous proof of the Bolzano–Weierstrass Theorem and its use to prove some of the difficult theorems that are usually assumed in a one-variable calculus course. We shall skip over some easier aspects of an undergraduate course in real analysis that fit logically at the end of this section. A list of such topics appears at the end of the section.

The system of real numbers $\mathbb{R}$ may be constructed out of the system of rational numbers $\mathbb{Q}$, and we take this construction as known. The formal definition is that a real number is a **cut** of rational numbers, i.e., a subset of rational numbers that is neither $\mathbb{Q}$ nor the empty set, has no largest element, and contains all rational numbers less than any rational that it contains. The idea of the construction is as follows: Each rational number q determines a cut q^*, namely the set of all rationals less than q. Under the identification of $\mathbb{Q}$ with a subset of $\mathbb{R}$, the cut defining a real number consists of all rational numbers less than the given real number.

The set of cuts gets a natural ordering, given by inclusion. In place of $\subseteq$, we write $\leq$. For any two cuts r and s, we have $r \leq s$ or $s \leq r$, and if both occur, then $r = s$. We can then define $<$, $\geq$, and $>$ in the expected way. The positive cuts r are those with $0^* < r$, and the negative cuts are those with $r < 0^*$.

Once cuts and their ordering are in place, one can go about defining the usual operations of arithmetic and proving that $\mathbb{R}$ with these operations satisfies the familiar associative, commutative, and distributive laws, and that these interact with inequalities in the usual ways. The definitions of addition and subtraction are easy: the sum or difference of two cuts is simply the set of sums or differences of the rationals from the respective cuts. For multiplication and reciprocals one has to take signs into account. For example, the product of two positive cuts consists of all products of positive rationals from the two cuts, as well as 0 and all negative rationals. After these definitions and the proofs of the usual arithmetic operations are complete, it is customary to write 0 and 1 in place of 0^* and 1^*.

An **upper bound** for a nonempty subset E of $\mathbb{R}$ is a real number M such that $x \leq M$ for all x in E. If the nonempty set E has an upper bound, we can take the cuts that E consists of and form their union. This turns out to be a cut, it is an upper bound for E, and it is $\leq$ all upper bounds for E. We can summarize this result as a theorem.

Theorem 1.1. Any nonempty subset E of $\mathbb{R}$ with an upper bound has a least upper bound.

The least upper bound is necessarily unique, and the notation for it is $\sup_{x \in E} x$ or $\sup \{x \mid x \in E\}$, "sup" being an abbreviation for the Latin word "supremum," the largest. Of course, the least upper bound for a set E with an upper bound need not be in E; for example, the supremum of the negative rationals is 0, which is not negative.

A **lower bound** for a nonempty set E of $\mathbb{R}$ is a real number m such that $x \geq m$ for all $x \in E$. If m is a lower bound for E, then $-m$ is an upper bound for the set $-E$ of negatives of members of E. Thus $-E$ has an upper bound, and Theorem 1.1 shows that it has a least upper bound $\sup_{x \in -E} x$. Then $-x$ is a greatest lower bound for E. This greatest lower bound is denoted by $\inf_{y \in E} y$ or $\inf\{y \mid y \in E\}$, "inf" being an abbreviation for "infimum." We can summarize as follows.

Corollary 1.2. Any nonempty subset E of $\mathbb{R}$ with a lower bound has a greatest lower bound.

A subset of $\mathbb{R}$ is said to be **bounded** if it has an upper bound and a lower bound. Let us introduce notation and terminology for **intervals** of $\mathbb{R}$, first treating the bounded ones.[1] Let a and b be real numbers with $a \leq b$. The **open interval** from a to b is the set $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$, the **closed interval** is the set $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$, and the **half-open intervals** are the sets $[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$ and $(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$. Each of the above intervals is indeed bounded, having a as a lower bound and b as an upper bound. These intervals are nonempty when $a < b$ or when the interval is $[a, b]$ with $a = b$, and in these cases the least upper bound is b and the greatest lower bound is a.

Open sets in $\mathbb{R}$ are defined to be arbitrary unions of open bounded intervals, and a **closed set** is any set whose complement in $\mathbb{R}$ is open. A set E is open if and only if for each $x \in E$, there is an open interval (a, b) such that $x \in (a, b) \subseteq E$. In this case we of course have $a < x < b$. If we put $\epsilon = \min\{x - a, b - x\}$, then we see that x lies in the subset $(x - \epsilon, x + \epsilon)$ of (a, b). The open interval $(x - \epsilon, x + \epsilon)$ equals $\{y \in \mathbb{R} \,\big|\, |y - x| < \epsilon\}$. Thus an open set in $\mathbb{R}$ is any set E such that for each $x \in E$, there is a number $\epsilon > 0$ such that $\{y \in \mathbb{R} \,\big|\, |y - x| < \epsilon\}$ lies in E. A **limit point** x of a subset F of $\mathbb{R}$ is a point of $\mathbb{R}$ such that any open interval containing x meets F in a point other than x. For example, the set $[a, b) \cup \{b + 1\}$ has $[a, b]$ as its set of limit points. A subset of $\mathbb{R}$ is closed if and only if it contains all its limit points.

Now let us turn to unbounded intervals. To provide notation for these, we shall make use of two symbols $+\infty$ and $-\infty$ that will shortly be defined to be "extended real numbers." If a is in $\mathbb{R}$, then the subsets $(a, +\infty) = \{x \in \mathbb{R} \mid a < x\}$, $(-\infty, a) = \{x \in \mathbb{R} \mid x < a\}$, $(-\infty, +\infty) = \mathbb{R}$, $[a, +\infty) = \{x \in \mathbb{R} \mid a \leq x\}$, and $(-\infty, a] = \{x \in \mathbb{R} \mid x \leq a\}$ are defined to be **intervals**, and they are all unbounded. The first three are open sets of $\mathbb{R}$ and are considered to be open

[1] Bounded intervals are called "finite intervals" by some authors.

intervals, while the last three are closed sets and are considered to be closed intervals. Specifically the middle set $\mathbb{R}$ is both open and closed.

One important consequence of Theorem 1.1 is the **archimedean property** of $\mathbb{R}$, as follows.

Corollary 1.3. If a and b are real numbers with $a > 0$, then there exists an integer n with $na > b$.

PROOF. If, on the contrary, $na \le b$ for all integers n, then b is an upper bound for the set of all na. Let M be the least upper bound of the set $\{na \mid n \text{ is an integer}\}$. Using that a is positive, we find that $a^{-1}M$ is a least upper bound for the integers. Thus $n \le a^{-1}M$ for all integers n, and there is no smaller upper bound. However, the smaller number $a^{-1}M - 1$ must be an upper bound, since saying $n \le a^{-1}M$ for all integers is the same as saying $n-1 \le a^{-1}M-1$ for all integers. We arrive at a contradiction, and we conclude that there is some integer n with $na > b$.

The archimedean property enables one to see, for example, that any two distinct real numbers have a rational number lying between them. We prove this consequence as Corollary 1.5 after isolating one step as Corollary 1.4.

Corollary 1.4. If c is a real number, then there exists an integer n such that $n \le c < n+1$.

PROOF. Corollary 1.3 with $a = 1$ and $b = c$ shows that there is an integer M with $M > c$, and Corollary 1.3 with $a = 1$ and $b = -c$ shows that there is an integer m with $m > -c$. Then $-m < c < M$, and it follows that there exists a greatest integer n with $n \le c$. This n must have the property that $c < n+1$, and the corollary follows.

Corollary 1.5. If x and y are real numbers with $x < y$, then there exists a rational number r with $x < r < y$.

PROOF. By Corollary 1.3 with $a = y - x$ and $b = 1$, there is an integer N such that $N(y-x) > 1$. This integer N has to be positive. Then $\frac{1}{N} < y - x$. By Corollary 1.4 with $c = Nx$, there exists an integer n with $n \le Nx < n+1$, hence with $\frac{n}{N} \le x < \frac{n+1}{N}$. Adding the inequalities $\frac{n}{N} \le x$ and $\frac{1}{N} < y - x$ yields $\frac{n+1}{N} < y$. Thus $x \le \frac{n}{N} < \frac{n+1}{N} < y$. Since $\frac{n}{N} < \frac{2n+1}{2N} < \frac{n+1}{N}$, the rational number $r = \frac{2n+1}{2N}$ has the required properties.

A **sequence** in a set S is a function from a certain kind of subset of integers into S. It will be assumed that the set of integers is nonempty, consists of consecutive integers, and contains no largest integer. In particular the domain of any sequence is infinite. Usually the set of integers is either all nonnegative integers or all positive integers. Sometimes the set of integers is all integers, and the sequence in this case is often called "doubly infinite." The value of a sequence f at the integer n is normally written f_n rather than $f(n)$, and the sequence itself may be denoted by an expression like $\{f_n\}_{n\geq 1}$, in which the outer subscript indicates the domain.

A **subsequence** of a sequence f with domain $\{m, m+1, \dots\}$ is a composition $f \circ n$, where f is a sequence and n is a sequence in the domain of f such that $n_k < n_{k+1}$ for all k. For example, if $\{a_n\}_{n\geq 1}$ is a sequence, then $\{a_{2k}\}_{k\geq 1}$ is the subsequence in which the function n is given by $n_k = 2k$. The domain of a subsequence, by our definition, is always infinite.

A sequence a_n in $\mathbb{R}$ is **convergent**, or convergent in $\mathbb{R}$, if there exists a real number a such that for each $\epsilon > 0$, there is an integer N with $|a_n - a| < \epsilon$ for all $n \geq N$. The number a is necessarily unique and is called the **limit** of the sequence. Depending on how much information about the sequence is unambiguous, we may write $\lim_{n\to\infty} a_n = a$ or $\lim_n a_n = a$ or $\lim a_n = a$ or $a_n \to a$. We also say a_n **tends** to a as n tends to **infinity** or ∞.

A sequence in $\mathbb{R}$ is called **monotone increasing** if $a_n \leq a_{n+1}$ for all n in the domain, **monotone decreasing** if $a_n \geq a_{n+1}$ for all n in the domain, **monotone** if it is monotone increasing or monotone decreasing.

Corollary 1.6. Any bounded monotone sequence in $\mathbb{R}$ converges. If the sequence is monotone increasing, then the limit is the least upper bound of the image in $\mathbb{R}$ of the sequence. If the sequence is monotone decreasing, the limit is the greatest lower bound of the image.

REMARK. Often it is Corollary 1.6, rather than the existence of least upper bounds, that is taken for granted in an elementary calculus course. The reason is that the statement of Corollary 1.6 tends for calculus students to be easier to understand than the statement of the least upper bound property. Problem 1 at the end of the chapter asks for a derivation of the least-upper-bound property from Corollary 1.6.

PROOF. Suppose that $\{a_n\}$ is monotone increasing and bounded. Let $a = \sup_n a_n$, the existence of the supremum being ensured by Theorem 1.1, and let $\epsilon > 0$ be given. If there were no integer N with $a_N > a - \epsilon$, then $a - \epsilon$ would be a smaller upper bound, contradiction. Thus such an N exists. For that N, $n \geq N$ implies $a - \epsilon < a_N \leq a_n \leq a < a + \epsilon$. Thus $n \geq N$ implies $|a_n - a| < \epsilon$. Since ϵ is arbitrary, $\lim_{n\to\infty} a_n = a$. If the given sequence $\{a_n\}$ is monotone decreasing, we argue similarly with $a = \inf_n a_n$.

In working with sup and inf, it will be quite convenient to use the notation $\sup_{x\in E} x$ even when E is nonempty but not bounded above, and to use the notation

$\inf_{x\in E} x$ even when E is nonempty but not bounded below. We introduce symbols $+\infty$ and $-\infty$, plus and minus **infinity**, for this purpose and extend the definitions of $\sup_{x\in E} x$ and $\inf_{x\in E} x$ to all nonempty subsets E of $\mathbb{R}$ by taking

$$\sup_{x\in E} x = +\infty \qquad \text{if } E \text{ has no upper bound,}$$
$$\inf_{x\in E} x = -\infty \qquad \text{if } E \text{ has no lower bound.}$$

To work effectively with these new pieces of notation, we shall enlarge $\mathbb{R}$ to a set $\mathbb{R}^*$ called the **extended real numbers** by defining

$$\mathbb{R}^* = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}.$$

An ordering on $\mathbb{R}^*$ is defined by taking $-\infty < r < +\infty$ for every member r of $\mathbb{R}$ and by retaining the usual ordering within $\mathbb{R}$. It is immediate from this definition that

$$\inf_{x\in E} x \le \sup_{x\in E} x$$

if E is any nonempty subset of $\mathbb{R}$. In fact, we can enlarge the definitions of $\inf_{x\in E} x$ and $\sup_{x\in E} x$ in obvious fashion to include the case that E is any nonempty subset of $\mathbb{R}^*$, and we still have $\inf \le \sup$. With the ordering in place, we can unambiguously speak of **open intervals** (a, b), **closed intervals** $[a, b]$, and **half-open intervals** $[a, b)$ and $(a, b]$ in $\mathbb{R}^*$ even if a or b is infinite. Under our definitions the intervals of $\mathbb{R}$ are the intervals of $\mathbb{R}^*$ that are subsets of $\mathbb{R}$, even if a or b is infinite. If no special mention is made whether an interval lies in $\mathbb{R}$ or $\mathbb{R}^*$, it is usually assumed to lie in $\mathbb{R}$.

The next step is to extend the operations of arithmetic to $\mathbb{R}^*$. It is important not to try to make such operations be everywhere defined, lest the distributive laws fail. Letting r denote any member of $\mathbb{R}$ and a and b be any members of $\mathbb{R}^*$, we make the following new definitions:

Multiplication:

$$r(+\infty) = (+\infty)r = \begin{cases} +\infty & \text{if } r > 0, \\ 0 & \text{if } r = 0, \\ -\infty & \text{if } r < 0, \end{cases}$$

$$r(-\infty) = (-\infty)r = \begin{cases} -\infty & \text{if } r > 0, \\ 0 & \text{if } r = 0, \\ +\infty & \text{if } r < 0, \end{cases}$$

$$(+\infty)(+\infty) = (-\infty)(-\infty) = +\infty,$$
$$(+\infty)(-\infty) = (-\infty)(+\infty) = -\infty.$$

Addition:
$$r + (+\infty) = (+\infty) + r = +\infty,$$
$$r + (-\infty) = (-\infty) + r = -\infty,$$
$$(+\infty) + (+\infty) = +\infty,$$
$$(-\infty) + (-\infty) = -\infty.$$

Subtraction: $a - b = a + (-b)$ whenever the right side is defined.

Division: $a/b = 0$ if $a \in \mathbb{R}$ and b is $\pm\infty$,

$a/b = b^{-1}a$ if $b \in \mathbb{R}$ with $b \neq 0$ and a is $\pm\infty$.

The only surprise in the list is that 0 times anything is 0. This definition will be important to us when we get to measure theory, starting in Chapter V.

It is now a simple matter to define convergence of a sequence in $\mathbb{R}^*$. The cases that need addressing are that the sequence is in $\mathbb{R}$ and that the limit is $+\infty$ or $-\infty$. We say that a sequence $\{a_n\}$ in $\mathbb{R}$ tends to $+\infty$ if for any positive number M, there exists an integer N such that $a_n \geq M$ for all $n \geq N$. The sequence tends to $-\infty$ if for any negative number $-M$, there exists an integer N such that $a_n \leq -M$ for all $n \geq N$. It is important to indicate whether convergence/divergence of a sequence is being discussed in $\mathbb{R}$ or in $\mathbb{R}^*$. The default setting is $\mathbb{R}$, in keeping with standard terminology in calculus. Thus, for example, we say that the sequence $\{n\}_{n\geq 1}$ diverges, but it converges in $\mathbb{R}^*$ (to $+\infty$).

With our new definitions every monotone sequence converges in $\mathbb{R}^*$.

For a sequence $\{a_n\}$ in $\mathbb{R}$ or even in $\mathbb{R}^*$, we now introduce members $\limsup_n a_n$ and $\liminf_n a_n$ of $\mathbb{R}^*$. These will always be defined, and thus we can apply the operations lim sup and lim inf to any sequence in $\mathbb{R}^*$. For the case of lim sup we define $b_n = \sup_{k\geq n} a_k$ as a sequence in $\mathbb{R}^*$. The sequence $\{b_n\}$ is monotone decreasing. Thus it converges to $\inf_n b_n$ in $\mathbb{R}^*$. We define[2]

$$\limsup_n a_n = \inf_n \sup_{k\geq n} a_k$$

as a member of $\mathbb{R}^*$, and we define

$$\liminf_n a_n = \sup_n \inf_{k\geq n} a_k$$

as a member of $\mathbb{R}^*$. Let us underscore that $\limsup a_n$ and $\liminf a_n$ always exist. However, one or both may be $\pm\infty$ even if a_n is in $\mathbb{R}$ for every n.

Proposition 1.7. The operations lim sup and lim inf on sequences $\{a_n\}$ and $\{b_n\}$ in $\mathbb{R}^*$ have the following properties:

(a) if $a_n \leq b_n$ for all n, then $\limsup a_n \leq \limsup b_n$ and $\liminf a_n \leq \liminf b_n$,

[2]The notation $\overline{\lim}$ was at one time used for lim sup, and $\underline{\lim}$ was used for lim inf.

(b) $\liminf a_n \le \limsup a_n$,

(c) $\{a_n\}$ has a subsequence converging in $\mathbb{R}^*$ to $\limsup a_n$ and another subsequence converging in $\mathbb{R}^*$ to $\liminf a_n$,

(d) $\limsup a_n$ is the supremum of all subsequential limits of $\{a_n\}$ in $\mathbb{R}^*$, and $\liminf_n$ is the infimum of all subsequential limits of $\{a_n\}$ in $\mathbb{R}^*$,

(e) if $\limsup a_n < +\infty$, then $\limsup a_n$ is the infimum of all extended real numbers a such that $a_n \ge a$ for only finitely many n, and if $\liminf a_n > -\infty$, then $\liminf a_n$ is the supremum of all extended real numbers a such that $a_n \le a$ for only finitely many n,

(f) the sequence $\{a_n\}$ in $\mathbb{R}^*$ converges in $\mathbb{R}^*$ if and only if $\liminf a_n = \limsup a_n$, and in this case the limit is the common value of $\liminf a_n$ and $\limsup a_n$.

REMARK. It is enough to prove the results about lim sup, since $\liminf a_n = -\limsup(-a_n)$.

PROOFS FOR lim sup.

(a) From $a_l \le b_l$ for all l, we have $a_l \le \sup_{k\ge n} b_k$ if $l \ge n$. Hence $\sup_{l\ge n} a_l \le \sup_{k\ge n} b_k$. Then (a) follows by taking the limit on n.

(b) This follows by taking the limit on n of the inequality $\inf_{k\ge n} a_k \le \sup_{k\ge n} a_k$.

(c) We divide matters into cases. The main case is that $a = \limsup a_n$ is in $\mathbb{R}$. Inductively, for each $l \ge 1$, choose $N \ge n_{l-1}$ such that $|\sup_{k\ge N} a_k - a| < l^{-1}$. Then choose $n_l > n_{l-1}$ such that $|a_{n_l} - \sup_{k\ge N} a_k| < l^{-1}$. Together these inequalities imply $|a_{n_l} - a| < 2l^{-1}$ for all l, and thus $\lim_{l\to\infty} a_{n_l} = a$. The second case is that $a = \limsup a_n$ equals $+\infty$. Since $\sup_{k\ge n} a_k$ is monotone decreasing in n, we must have $\sup_{k\ge n} a_k = +\infty$ for all n. Inductively for $l \ge 1$, we can choose $n_l > n_{l-1}$ such that $a_{n_l} \ge l$. Then $\lim_{l\to\infty} a_{n_l} = +\infty$. The third case is that $a = \limsup a_n$ equals $-\infty$. The sequence $b_n = \sup_{k\ge n} a_k$ is monotone decreasing to $-\infty$. Inductively for $l \ge 1$, choose $n_l > n_{l-1}$ such that $b_{n_l} \le -l$. Then $a_{n_l} \le b_{n_l} \le -l$, and $\lim_{l\to\infty} a_{n_l} = -\infty$.

(d) By (c), $\limsup a_n$ is one subsequential limit. Let $a = \lim_{k\to\infty} a_{n_k}$ be another subsequential limit. Put $b_n = \sup_{l\ge n} a_l$. Then $\{b_n\}$ converges to $\limsup a_n$ in $\mathbb{R}^*$, and the same thing is true of every subsequence. Since $a_{n_k} \le \sup_{l\ge n_k} a_l = b_{n_k}$ for all k, we can let k tend to infinity and obtain $a = \lim_{k\to\infty} a_{n_k} \le \lim_{k\to\infty} b_{n_k} = \limsup a_n$.

(e) Since $\limsup a_n < +\infty$, we have $\sup_{k\ge n} a_k < +\infty$ for n greater than or equal to some N. For this N and any $a > \sup_{k\ge N} a_k$, we then have $a_n \ge a$ only finitely often. Thus there exists $a \in \mathbb{R}$ such that $a_n \ge a$ for only finitely many n. On the other hand, if a' is a real number $< \limsup a_n$, then (c) shows that $a_n \ge a'$ for infinitely many n. Hence

$$\limsup a_n \le \inf\{a \mid a_n \ge a \text{ for only finitely many } a\}.$$

Arguing by contradiction, suppose that $<$ holds in this inequality, and let a'' be a real number strictly in between the two sides of the inequality. Then $\sup_{k\geq n} a_k < a''$ for n large enough, and so $a_n \geq a''$ only finitely often. But then a'' is in the set

$$\{a \mid a_n \geq a \text{ for only finitely many } a\},$$

and the statement that a'' is less than the infimum of this set gives a contradiction.

(f) If $\{a_n\}$ converges in $\mathbb{R}^*$, then (c) forces $\liminf a_n = \limsup a_n$. Conversely suppose $\liminf a_n = \limsup a_n$, and let a be the common value of $\liminf a_n$ and $\limsup a_n$. The main case is that a is in $\mathbb{R}$. Let $\epsilon > 0$ be given. By (e), $a_n \geq a+\epsilon$ only finitely often, and $a_n \leq a - \epsilon$ only finitely often. Thus $|a_n - a| < \epsilon$ for all n sufficiently large. In other words, $\lim a_n = a$ as asserted. The other cases are that $a = +\infty$ or $a = -\infty$, and they are completely analogous to each other. Suppose for definiteness that $a = +\infty$. Since $\liminf a_n = +\infty$, the monotone increasing sequence $b_n = \inf_{k\geq n} a_k$ converges in $\mathbb{R}^*$ to $+\infty$. Given M, choose N such that $b_n \geq M$ for $n \geq N$. Then also $a_n \geq M$ for $n \geq N$, and a_n converges in $\mathbb{R}^*$ to $+\infty$. This completes the proof.

With Proposition 1.7 as a tool, we can now prove the Bolzano–Weierstrass Theorem. The remainder of the section will consist of applications of this theorem, showing that Cauchy sequences in $\mathbb{R}$ converge in $\mathbb{R}$, that continuous functions on closed bounded intervals of $\mathbb{R}$ are uniformly continuous, that continuous functions on closed bounded intervals are bounded and assume their maximum and minimum values, and that continuous functions on closed intervals take on all intermediate values.

Theorem 1.8 (Bolzano–Weierstrass). Every bounded sequence in $\mathbb{R}$ has a convergent subsequence with limit in $\mathbb{R}$.

PROOF. If the given bounded sequence is $\{a_n\}$, form the subsequence noted in Proposition 1.7c that converges in $\mathbb{R}^*$ to $\limsup a_n$. All quantities arising in the formation of $\limsup a_n$ are in $\mathbb{R}$, since $\{a_n\}$ is bounded, and thus the limit is in $\mathbb{R}$.

A sequence $\{a_n\}$ in $\mathbb{R}$ is called a **Cauchy sequence** if for any $\epsilon > 0$, there exists an N such that $|a_n - a_m| < \epsilon$ for all n and m that are $\geq N$.

EXAMPLE. Every convergent sequence in $\mathbb{R}$ with limit in $\mathbb{R}$ is Cauchy. In fact, let $a = \lim a_n$, and let $\epsilon > 0$ be given. Choose N such that $n \geq N$ implies $|a_n - a| < \epsilon$. Then $n, m \geq N$ implies

$$|a_n - a_m| \leq |a_n - a| + |a - a_m| < \epsilon + \epsilon = 2\epsilon.$$

Hence the sequence is Cauchy.

In the above example and elsewhere in this book, we allow ourselves the luxury of having our final bound come out as a fixed multiple $M\epsilon$ of ϵ, rather than ϵ itself. Strictly speaking, we should have introduced $\epsilon' = \epsilon/M$ and aimed for ϵ' rather than ϵ. Then our final bound would have been $M\epsilon' = \epsilon$. Since the technique for adjusting a proof in this way is always the same, we shall not add these extra steps in the future unless there would otherwise be a possibility of confusion.

This convention suggests a handy piece of terminology—that a proof as in the above example, in which $M = 2$, is a "2ϵ proof." That name conveys a great deal of information about the proof, saying that one should expect two contributions to the final estimate and that the final bound will be 2ϵ.

Theorem 1.9 (Cauchy criterion). Every Cauchy sequence in $\mathbb{R}$ converges to a limit in $\mathbb{R}$.

PROOF. Let $\{a_n\}$ be Cauchy in $\mathbb{R}$. First let us see that $\{a_n\}$ is bounded. In fact, for $\epsilon = 1$, choose N such that $n, m \geq N$ implies $|a_n - a_m| < 1$. Then $|a_m| \leq |a_N| + 1$ for $m \geq N$, and $M = \max\{|a_1|, \ldots, |a_{N-1}|, |a_N| + 1\}$ is a common bound for all $|a_n|$.

Since $\{a_n\}$ is bounded, it has a convergent subsequence $\{a_{n_k}\}$, say with limit a, by the Bolzano–Weierstrass Theorem. The subsequential limit has to satisfy $|a| \leq M$ within $\mathbb{R}^*$, and thus a is in $\mathbb{R}$.

Finally let us see that $\lim a_n = a$. In fact, if $\epsilon > 0$ is given, choose N such that $n_k \geq N$ implies $|a_{n_k} - a| < \epsilon$. Also, choose $N' \geq N$ such that $n, m \geq N'$ implies $|a_n - a_m| < \epsilon$. If $n \geq N'$, then any $n_k \geq N'$ has $|a_n - a_{n_k}| < \epsilon$, and hence

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \epsilon + \epsilon = 2\epsilon.$$

This completes the proof.

Let f be a function with domain an interval and with range in $\mathbb{R}$. The interval is allowed to be unbounded, but it is required to be a subset of $\mathbb{R}$. We say that f is **continuous at** a point x_0 of the domain of f within $\mathbb{R}$ if for each $\epsilon > 0$, there is some $\delta > 0$ such that all x in the domain of f that satisfy $|x - x_0| < \delta$ have $|f(x) - f(x_0)| < \epsilon$. This notion is sometimes abbreviated as $\lim_{x \to x_0} f(x) = f(x_0)$. Alternatively, one may say that $f(x)$ tends to $f(x_0)$ as x tends to x_0, and one may write $f(x) \to f(x_0)$ as $x \to x_0$.

A mathematically equivalent definition is that f is continuous at x_0 if whenever a sequence has $x_n \to x_0$ within the domain interval, then $f(x_n) \to f(x_0)$. This latter version of continuity will be shown in Section II.4 to be equivalent to the former version, given in terms of continuous limits, in greater generality than just for $\mathbb{R}$, and thus we shall not stop to prove the equivalence now. We say that f is **continuous** if it is continuous at all points of its domain.

We say that the a function f as above is **uniformly continuous** on its domain if for any $\epsilon > 0$, there is some $\delta > 0$ such that $|f(x) - f(x_0)| < \epsilon$ whenever x and x_0 are in the domain interval and $|x - x_0| < \delta$. (In other words, the condition for the continuity to be uniform is that δ can always be chosen independently of x_0.)

EXAMPLE. The function $f(x) = x^2$ is continuous on $(-\infty, +\infty)$, but it is not uniformly continuous. In fact, it is not uniformly continuous on $[1, +\infty)$. Assuming the contrary, choose δ for $\epsilon = 1$. Then we must have $\left|(x+\frac{\delta}{2})^2 - x^2\right| < 1$ for all $x \geq 1$. But $\left|(x + \frac{\delta}{2})^2 - x^2\right| = \delta x + \frac{\delta^2}{4} \geq \delta x$, and this is ≥ 1 for $x \geq \delta^{-1}$.

Theorem 1.10. A continuous function f from a closed bounded interval $[a, b]$ into $\mathbb{R}$ is uniformly continuous.

PROOF. Fix $\epsilon > 0$. For x_0 in the domain of f, the continuity of f at x_0 means that it makes sense to define

$$\delta_{x_0}(\epsilon) = \min\left\{1,\ \sup\left\{\delta' > 0 \,\middle|\, \begin{array}{l} |x - x_0| < \delta' \text{ and } x \text{ in the domain} \\ \text{of } f \text{ imply } |f(x) - f(x_0)| < \epsilon \end{array}\right\}\right\}.$$

If $|x - x_0| < \delta_{x_0}(\epsilon)$, then $|f(x) - f(x_0)| < \epsilon$. Put $\delta(\epsilon) = \inf_{x_0 \in [a,b]} \delta_{x_0}(\epsilon)$. Let us see that it is enough to prove that $\delta(\epsilon) > 0$. If x and y are in $[a, b]$ with $|x - y| < \delta(\epsilon)$, then $|x - y| < \delta(\epsilon) \leq \delta_y(\epsilon)$. Hence $|f(x) - f(y)| < \epsilon$ as required.

Thus we are to prove that $\delta(\epsilon) > 0$. If $\delta(\epsilon) = 0$, then, for each integer $n > 0$, we can choose x_n such that $\delta_{x_n}(\epsilon) < \frac{1}{n}$. By the Bolzano–Weierstrass Theorem, there is a convergent subsequence, say with $x_{n_k} \to x'$. Along this subsequence, $\delta_{x_{n_k}}(\epsilon) \to 0$. Fix k large enough so that $|x_{n_k} - x'| < \frac{1}{2}\delta_{x'}(\frac{\epsilon}{2})$. Then $|f(x_{n_k}) - f(x')| < \frac{\epsilon}{2}$. Also, $|x - x_{n_k}| < \frac{1}{2}\delta_{x'}(\frac{\epsilon}{2})$ implies

$$|x - x'| \leq |x - x_{n_k}| + |x_{n_k} - x'| < \tfrac{1}{2}\delta_{x'}(\tfrac{\epsilon}{2}) + \tfrac{1}{2}\delta_{x'}(\tfrac{\epsilon}{2}) = \delta_{x'}(\tfrac{\epsilon}{2}),$$

so that $|f(x) - f(x')| < \frac{\epsilon}{2}$ and

$$|f(x_{n_k}) - f(x)| \leq |f(x_{n_k}) - f(x')| + |f(x') - f(x)| < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon.$$

Consequently our arbitrary large fixed k has $\delta_{x_{n_k}} \geq \frac{1}{2}\delta_{x'}(\frac{\epsilon}{2})$, and the sequence $\{\delta_{x_{n_k}}(\epsilon)\}$ cannot be tending to 0.

Theorem 1.11. A continuous function f from a closed bounded interval $[a, b]$ into $\mathbb{R}$ is bounded and takes on maximum and minimum values.

PROOF. Let $c = \sup_{x \in [a,b]} f(x)$ in $\mathbb{R}^*$. Choose a sequence x_n in $[a, b]$ with $f(x_n)$ increasing to c. By the Bolzano–Weierstrass Theorem, $\{x_n\}$ has a convergent subsequence, say $x_{n_k} \to x'$. By continuity, $f(x_{n_k}) \to f(x')$. Then $f(x') = c$, and c is a finite maximum. The proof for a finite minimum is similar.

Theorem 1.12 (Intermediate Value Theorem). Let $a < b$ be real numbers, and let $f : [a, b] \to \mathbb{R}$ be continuous. Then f, in the interval $[a, b]$, takes on all values between $f(a)$ and $f(b)$.

REMARK. The proof below, which uses the Bolzano–Weierstrass Theorem, does not make absolutely clear what aspects of the structure of $\mathbb{R}$ are essential to the argument. A conceptually clearer proof will be given in Section II.8 and will bring out that the essential property of the interval $[a, b]$ is its "connectedness" in a sense to be defined in that section.

PROOF. Let $f(a) = \alpha$ and $f(b) = \beta$, and let γ be between α and β. We may assume that γ is in fact strictly between α and β. Possibly by replacing f by $-f$, we may assume that also $\alpha < \beta$. Let

$$A = \{x \in [a, b] \mid f(x) \leq \gamma\} \qquad \text{and} \qquad B = \{x \in [a, b] \mid f(x) \geq \gamma\}.$$

These sets are nonempty, since a is in A and b is in B, and f is bounded as a result of Theorem 1.11. Thus the numbers $\gamma_1 = \sup\{f(x) \mid x \in A\}$ and $\gamma_2 = \inf\{f(x) \mid x \in B\}$ are well defined and have $\gamma_1 \leq \gamma \leq \gamma_2$.

If $\gamma_1 = \gamma$, then we can find a sequence $\{x_n\}$ in A such that $f(x_n)$ converges to γ. Using the Bolzano–Weierstrass Theorem, we can find a convergent subsequence $\{x_{n_k}\}$ of $\{x_n\}$, say with limit x_0. By continuity of f, $\{f(x_{n_k})\}$ converges to $f(x_0)$. Then $f(x_0) = \gamma_1 = \gamma$, and we are done. Arguing by contradiction, we may therefore assume that $\gamma_1 < \gamma$. Similarly we may assume that $\gamma < \gamma_2$, but we do not need to do so.

Let $\epsilon = \gamma_2 - \gamma_1$, and choose, by Theorem 1.10 and uniform continuity, $\delta > 0$ such that $|x_1 - x_2| < \delta$ implies $|f(x_1) - f(x_2)| < \epsilon$ whenever x_1 and x_2 both lie in $[a, b]$. Then choose an integer n such that $2^{-n}(b - a) < \delta$, and consider the value of f at the points $p_k = a + k2^{-n}(b - a)$ for $0 \leq k \leq 2^{-n}$. Since $p_{k+1} - p_k = 2^{-n}(b - a) < \delta$, we have $|f(p_{k+1}) - f(p_k)| < \epsilon = \gamma_2 - \gamma_1$. Consequently if $f(p_k) \leq \gamma_1$, then

$$f(p_{k+1}) \leq f(p_k) + |f(p_{k+1}) - f(p_k)| < \gamma_1 + (\gamma_2 - \gamma_1) = \gamma_2,$$

and hence $f(p_{k+1}) \leq \gamma_1$. Now $f(p_0) = f(a) = \alpha \leq \gamma_1$. Thus induction shows that $f(p_k) \leq \gamma_1$ for all $k \leq 2^{-n}$. However, for $k = 2^n$, we have $p_{2^{-n}} = b$, and $f(b) = \beta \geq \gamma > \gamma_1$, and we have arrived at a contradiction.

Further topics. Here a number of other topics of an undergraduate course in real-variable theory fit well logically. Among these are countable vs. uncountable sets, infinite series and tests for their convergence, the fact that every rearrangement of an infinite series of positive terms has the same sum, special sequences, derivatives, the Mean Value Theorem as in Section A2 of the appendix, and continuity and differentiability of inverse functions as in Section A3 of the appendix. We shall not stop here to review these topics, which are treated in many books. One such book is Rudin's *Principles of Mathematical Analysis*, the relevant chapters being 1 to 5. In Chapter 2 of that book, only the first few pages are needed; they are the ones where countable and uncountable sets are discussed.

2. Interchange of Limits

Let $\{b_{ij}\}$ be a doubly indexed sequence of real numbers. It is natural to ask for the extent to which

$$\lim_i \lim_j b_{ij} = \lim_j \lim_i b_{ij},$$

more specifically to ask how to tell, in an expression involving iterated limits, whether we can interchange the order of the two limit operations. We can view matters conveniently in terms of an infinite matrix

$$\begin{pmatrix} b_{11} & b_{12} & \cdots \\ b_{21} & b_{22} & \\ \vdots & & \ddots \end{pmatrix}.$$

The left-hand iterated limit, namely $\lim_i \lim_j b_{ij}$, is obtained by forming the limit of each row, assembling the results, and then taking the limit of the row limits down through the rows. The right-hand iterated limit, namely $\lim_j \lim_i b_{ij}$, is obtained by forming the limit of each column, assembling the results, and then taking the limit of the column limits through the columns. If we use the particular infinite matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots \\ 0 & 1 & 1 & 1 & \cdots \\ 0 & 0 & 1 & 1 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & & & & \ddots \end{pmatrix},$$

then we see that the first iterated limit depends only on the part of the matrix above the main diagonal, while the second iterated limit depends only on the part of the matrix below the main diagonal. Thus the two iterated limits in general have no reason at all to be related. In the specific matrix that we have just considered, they are 1 and 0, respectively. Let us consider some examples along the same lines but with an analytic flavor.

EXAMPLES.

(1) Let $b_{ij} = \dfrac{j}{i+j}$. Then $\lim_i \lim_j b_{ij} = 1$, while $\lim_j \lim_i b_{ij} = 0$.

(2) Let F_n be a continuous real-valued function on $\mathbb{R}$, and suppose that $F(x) = \lim F_n(x)$ exists for every x. Is F continuous? This is the same kind of question. It asks whether $\lim_{t\to x} F(t) \overset{?}{=} F(x)$, hence whether

$$\lim_{t\to x} \lim_{n\to\infty} F_n(t) \overset{?}{=} \lim_{n\to\infty} \lim_{t\to x} F_n(t).$$

If we take $f_k(x) = \dfrac{x^2}{(1+x^2)^k}$ for $k \geq 0$ and define $F_n(x) = \sum_{k=0}^n f_k(x)$, then each F_n is continuous. The sequence of functions $\{F_n\}$ has a pointwise limit $F(x) = \sum_{k=0}^\infty \dfrac{x^2}{(1+x^2)^k}$. The series is a geometric series, and we can easily calculate explicitly the partial sums and the limit function. The latter is

$$F(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 + x^2 & \text{if } x \neq 0. \end{cases}$$

It is apparent that the limit function is discontinuous.

(3) Let $\{f_n\}$ be a sequence of differentiable functions, and suppose that $f(x) = \lim f_n(x)$ exists for every x and is differentiable. Is $\lim f_n'(x) = f'(x)$? This question comes down to whether

$$\lim_{n\to\infty} \lim_{t\to x} \frac{f_n(t) - f_n(x)}{t - x} \stackrel{?}{=} \lim_{t\to x} \lim_{n\to\infty} \frac{f_n(t) - f_n(x)}{t - x}.$$

An example where the answer is negative uses the sine and cosine functions, which are undefined in the rigorous development until Section 7 on power series. The example has $f_n(x) = \dfrac{\sin nx}{\sqrt{n}}$ for $n \geq 1$. Then $\lim_n f_n(x) = 0$, so that $f(x) = 0$ and $f'(x) = 0$. Also, $f_n'(x) = \sqrt{n} \cos nx$, so that $f_n'(0) = \sqrt{n}$ does not tend to $0 = f'(0)$.

Yet we know many examples from calculus where an interchange of limits is valid. For example, in calculus of two variables, the first partial derivatives of nice functions—polynomials, for example—can be computed in either order with the same result, and double integrals of continuous functions over a rectangle can be calculated as iterated integrals in either order with the same result. Positive theorems about interchanging limits are usually based on some kind of uniform behavior, in a sense that we take up in the next section. A number of positive results of this kind ultimately come down to the following general theorem about doubly indexed sequences that are monotone increasing in each variable. In Section 3 we shall examine the mechanism of this theorem closely: the proof shows that the equality in question is $\sup_i \sup_j b_{ij} = \sup_j \sup_i b_{ij}$ and that it holds because both sides equal $\sup_{i,j} b_{ij}$.

Theorem 1.13. Let b_{ij} be members of $\mathbb{R}^*$ that are ≥ 0 for all i and j. Suppose that b_{ij} is monotone increasing in i, for each j, and is monotone increasing in j, for each i. Then

$$\lim_i \lim_j b_{ij} = \lim_j \lim_i b_{ij},$$

with all the indicated limits existing in $\mathbb{R}^*$.

PROOF. Put $L_i = \lim_j b_{ij}$ and $L'_j = \lim_i b_{ij}$. These limits exist in $\mathbb{R}^*$, since the sequences in question are monotone. Then $L_i \le L_{i+1}$ and $L'_j \le L'_{j+1}$, and thus

$$L = \lim_i L_i \qquad \text{and} \qquad L' = \lim_j L'_j$$

both exist in $\mathbb{R}^*$. Arguing by contradiction, suppose that $L < L'$. Then we can choose j_0 such that $L'_{j_0} > L$. Since $L'_{j_0} = \lim_i b_{ij_0}$, we can choose i_0 such that $b_{i_0 j_0} > L$. Then we have $L < b_{i_0 j_0} \le L_{i_0} \le L$, contradiction. Similarly the assumption $L' < L$ leads to a contradiction. We conclude that $L = L'$.

Corollary 1.14. If a_{lj} are members of $\mathbb{R}^*$ that are ≥ 0 and are monotone increasing in j for each l, then

$$\lim_j \sum_l a_{lj} = \sum_l \lim_j a_{lj}$$

in $\mathbb{R}^*$, the limits existing.

REMARK. This result will be generalized by the Monotone Convergence Theorem when we study abstract measure theory in Chapter V.

PROOF. Put $b_{ij} = \sum_{l=1}^{i} a_{lj}$ in Theorem 1.13.

Corollary 1.15. If c_{ij} are members of $\mathbb{R}^*$ that are ≥ 0 for all i and j, then

$$\sum_i \sum_j c_{ij} = \sum_j \sum_i c_{ij}$$

in $\mathbb{R}^*$, the limits existing.

REMARK. This result will be generalized by Fubini's Theorem when we study abstract measure theory in Chapter V.

PROOF. This follows from Corollary 1.14.

3. Uniform Convergence

Let us examine more closely what is happening in the proof of Theorem 1.13, in which it is proved that iterated limits can be interchanged under certain hypotheses of monotonicity. One of the iterated limits is $L = \lim_i \lim_j b_{ij}$, and the claim is that L is approached as i and j tend to infinity jointly. In terms of a matrix whose entries are the various b_{ij}'s, the pictorial assertion is that all the terms far down and to the right are close to L:

$$\begin{pmatrix} \cdots & \cdots \\ \cdots & \begin{array}{|l}\hline \text{All terms here} \\ \text{are close to } L \end{array} \end{pmatrix}.$$

To see this claim, let us choose a row limit L_{i_0} that is close to L and then take an entry $b_{i_0 j_0}$ that is close to L_{i_0}. Then $b_{i_0 j_0}$ is close to L, and all terms down and to the right from there are even closer because of the hypothesis of monotonicity.

To relate this behavior to something uniform, suppose that $L < +\infty$, and let some $\epsilon > 0$ be given. We have just seen that we can arrange to have $|L - b_{ij}| < \epsilon$ whenever $i \geq i_0$ and $j \geq j_0$. Then $|L_i - b_{ij}| < \epsilon$ whenever $i \geq i_0$, provided $j \geq j_0$. Also, we have $\lim_j b_{ij} = L_i$ for $i = 1, 2, \dots, i_0 - 1$. Thus $|L_i - b_{ij}| < \epsilon$ for all i, provided $j \geq j_0'$, where j_0' is some larger index than j_0. This is the notion of uniform convergence that we shall define precisely in a moment: an expression with a parameter (j in our case) has a limit (on the variable i in our case) with an estimate independent of the parameter. We can visualize matters as in the following matrix:

$$\begin{array}{cc} & \begin{array}{cc} j & j_0' \end{array} \\ i & \begin{pmatrix} \cdots & \begin{array}{|l} \text{All terms here} \\ \text{are close to } L_i \\ \text{on all rows.} \end{array} \end{pmatrix} \end{array}.$$

The vertical dividing line occurs when the column index j is equal to j_0', and all terms to the right of this line are close to their respective row limits L_i.

Let us see the effect of this situation on the problem of interchange of limits. The above diagram forces all the terms in the shaded part of $\begin{pmatrix} \cdots & \cdots \\ \cdots & \begin{array}{|l}\hline ////// \end{array} \end{pmatrix}$ to be close to one number if $\lim L_i$ exists, i.e., if the row limits are tending to a limit. If the other iterated limit exists, then it must be this same number. Thus the interchange of limits is valid under these circumstances.

Actually, we can get by with less. If, in the displayed diagram above, we assume that all the column limits L_j' exist, then it appears that all the column limits with $j \geq j_0'$ have to be close to the L_i's. From this we can deduce that the column limits have a limit L' and that the row limits L_i must tend to the limit of the column limits. In other words, the convergence of the rows in a suitable uniform fashion and the convergence of the columns together imply that both iterated limits exist and they are equal. We shall state this result rigorously as Proposition 1.16, which will become a prototype for applications later in this section.

Let S be a nonempty set, and let f and f_n, for integers $n \geq 1$, be functions from S to $\mathbb{R}$. We say that $f_n(x)$ **converges** to $f(x)$ **uniformly** for x in S if for any $\epsilon > 0$, there is an integer N such that $n \geq N$ implies $|f_n(x) - f(x)| < \epsilon$ for all x in S. It is equivalent to say that $\sup_{x \in E} |f_n(x) - f(x)|$ tends to 0 as n tends to infinity.

Proposition 1.16. Let b_{ij} be real numbers for $i \geq 1$ and $j \geq 1$. Suppose that

(i) $L_i = \lim_j b_{ij}$ exists in $\mathbb{R}$ uniformly in i, and
(ii) $L'_j = \lim_i b_{ij}$ exists in $\mathbb{R}$ for each j.

Then

(a) $L = \lim_i L_i$ exists in $\mathbb{R}$,
(b) $L' = \lim_j L'_j$ exists in $\mathbb{R}$,
(c) $L = L'$,
(d) the double limit on i and j of b_{ij} exists and equals the common value of the iterated limits L and L', i.e., for each $\epsilon > 0$, there exist i_0 and j_0 such that $|b_{ij} - L| < \epsilon$ whenever $i \geq i_0$ and $j \geq j_0$,
(e) $L'_j = \lim_i b_{ij}$ exists in $\mathbb{R}$ uniformly in j.

REMARK. In applications we shall sometimes have additional information, typically the validity of (a) or (b). According to the statement of the proposition, however, the conclusions are valid without taking this extra information as an additional hypothesis.

PROOF. Let $\epsilon > 0$ be given. By (i), choose j_0 such that $|b_{ij} - L_i| < \epsilon$ for all i whenever $j \geq j_0$. With $j \geq j_0$ fixed, (ii) says that $|b_{ij} - L'_j| < \epsilon$ whenever i is $\geq$ some $i_0 = i_0(j)$. For $j \geq j_0$ and $i \geq i_0(j)$, we then have

$$|L_i - L'_j| \leq |L_i - b_{ij}| + |b_{ij} - L'_j| < \epsilon + \epsilon = 2\epsilon.$$

If $j' \geq j_0$ and $i \geq i_0(j')$, we similarly have $|L_i - L'_{j'}| < 2\epsilon$. Hence if $j \geq j_0$, $j' \geq j_0$, and $i \geq \max\{i_0(j), i_0(j')\}$, then

$$|L'_j - L'_{j'}| \leq |L'_j - L_i| + |L_i - L'_{j'}| < 2\epsilon + 2\epsilon = 4\epsilon.$$

In other words, $\{L'_j\}$ is a Cauchy sequence. By Theorem 1.9, $L' = \lim_j L'_j$ exists in $\mathbb{R}$. This proves (b).

Passing to the limit in our inequality, we have $|L'_j - L'| \leq 4\epsilon$ when $j \geq j_0$ and in particular when $j = j_0$. If $i \geq i_0(j_0)$, then we saw that $|b_{ij_0} - L_i| < \epsilon$ and $|b_{ij_0} - L'_{j_0}| < \epsilon$. Hence $i \geq i_0(j_0)$ implies

$$|L_i - L'| \leq |L_i - b_{ij_0}| + |b_{ij_0} - L'_{j_0}| + |L'_{j_0} - L'| < \epsilon + \epsilon + 4\epsilon = 6\epsilon.$$

Since ϵ is arbitrary, $L = \lim_i L_i$ exists and equals L'. This proves (a) and (c).

Since $\lim_i L_i = L$, choose i_1 such that $|L_i - L| < \epsilon$ whenever $i \geq i_1$. If $i \geq i_1$ and $j \geq j_0$, we then have

$$|b_{ij} - L| \leq |b_{ij} - L_i| + |L_i - L| < \epsilon + \epsilon = 2\epsilon.$$

This proves (d).

Let i_1 and j_0 be as in the previous paragraph. We have seen that $|L'_j - L'_{j'}| < 4\epsilon$ for $j \geq j_0$. By (b), $|L'_j - L'| \leq 4\epsilon$ whenever $j \geq j_0$. Hence (c) and the inequality of the previous paragraph give

$$|b_{ij} - L'_j| \leq |b_{ij} - L| + |L - L'| + |L' - L'_j| < 2\epsilon + 0 + 4\epsilon = 6\epsilon$$

whenever $i \geq i_1$ and $j \geq j_0$. By (b), choose $j_1 \geq j_0$ such that $|b_{ij} - L'_j| < 6\epsilon$ whenever $i \in \{1, \ldots, i_1 - 1\}$ and $j \geq j_1$. Then $j \geq j_1$ implies $|b_{ij} - L'_j| < 6\epsilon$ for all i whenever $j \geq j_1$. This proves (e).

In checking for uniform convergence, we often do not have access to explicit expressions for limiting values. One device for dealing with the problem is a uniform version of the Cauchy criterion. Let S be a nonempty set, and let $\{f_n\}_{n\geq 1}$ be a sequence of functions from S to $\mathbb{R}$. We say that $\{f_n(x)\}$ is **uniformly Cauchy** for $x \in S$ if for any $\epsilon > 0$, there is an integer N such that $n \geq N$ and $m \geq N$ together imply $|f_n(x) - f_m(x)| < \epsilon$ for all x in S.

Proposition 1.17 (uniform Cauchy criterion). A sequence $\{f_n\}$ of functions from a nonempty set S to $\mathbb{R}$ is uniformly Cauchy if and only if it is uniformly convergent.

PROOF. If $\{f_n\}$ is uniformly convergent to f, we use a 2ϵ argument, just as in the example before Theorem 1.9: Given $\epsilon > 0$, choose N such that $n \geq N$ implies $|f_n(x) - f(x)| < \epsilon$. Then $n \geq N$ and $m \geq N$ together imply

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \epsilon + \epsilon = 2\epsilon.$$

Thus $\{f_n\}$ is uniformly Cauchy.

Conversely suppose that $\{f_n\}$ is uniformly Cauchy. Then $\{f_n(x)\}$ is Cauchy for each x. Theorem 1.9 therefore shows that there exists a function $f : S \to \mathbb{R}$ such that $\lim_n f_n(x) = f(x)$ for each x. We prove that the convergence is uniform. Given $\epsilon > 0$, choose N, as is possible since $\{f_n\}$ is uniformly Cauchy, such that $n \geq N$ and $m \geq N$ together imply $|f_n(x) - f_m(x)| < \epsilon$. Letting m tend to ∞ shows that $|f_n(x) - f(x)| \leq \epsilon$ for $n \geq N$. Hence $\lim_n f_n(x) = f(x)$ uniformly for x in S.

In practice, uniform convergence often arises with infinite series of functions, and then the definition and results about uniform convergence are to be applied to the sequence of partial sums. If the series is $\sum_{k=1}^{\infty} a_k(x)$, one wants $\big|\sum_{k=m}^{n} a_k(x)\big|$ to be small for all m and n sufficiently large. Some of the standard tests for convergence of series of numbers yield tests for uniform convergence of series of functions just by introducing a parameter and ensuring that the estimates do not depend on the parameter. We give two clear-cut examples. One is the uniform **alternating series test** or **Leibniz test**, given in Corollary 1.18. A generalization is the handy test given in Corollary 1.19.

Corollary 1.18. If for each x in a nonempty set S, $\{a_n(x)\}_{n\geq 1}$ is a monotone decreasing sequence of nonnegative real numbers such that $\lim_n a_n(x) = 0$ uniformly in x, then $\sum_{n=1}^{\infty} (-1)^n a_n(x)$ converges uniformly.

PROOF. The hypotheses are such that $\big|\sum_{k=m}^{n} (-1)^k a_k(x)\big| \leq \sup_x |a_m(x)|$ whenever $n \geq m$, and the uniform convergence is immediate from the uniform Cauchy criterion.

Corollary 1.19. If for each x in a nonempty set S, $\{a_n(x)\}_{n\geq 1}$ is a monotone decreasing sequence of nonnegative real numbers such that $\lim_n a_n(x) = 0$ uniformly in x and if $\{b_n(x)\}_{n\geq 1}$ is a sequence of real-valued functions on S whose partial sums $B_n(x) = \sum_{k=1}^{n} b_k(x)$ have $|B_n(x)| \leq M$ for some M and all n and x, then $\sum_{n=1}^{\infty} a_n(x) b_n(x)$ converges uniformly.

PROOF. If $n \geq m$, summation by parts gives

$$\sum_{k=m}^{n} a_k(x) b_k(x) = \sum_{k=m}^{n-1} B_k(x)(a_k(x) - a_{k+1}(x)) + B_n(x) a_n(x) - B_{m-1}(x) a_m(x).$$

Let $\epsilon > 0$ be given, and choose N such that $a_k(x) \leq \epsilon$ for all x whenever $k \geq N$. If $n \geq m \geq N$, then

$$\begin{aligned}
\Big|\sum_{k=m}^{n} a_k(x) b_k(x)\Big| &\leq \sum_{k=m}^{n-1} |B_k(x)|(a_k(x) - a_{k+1}(x)) + M\epsilon + M\epsilon \\
&\leq M \sum_{k=m}^{n-1} (a_k(x) - a_{k+1}(x)) + 2M\epsilon \\
&\leq M a_m(x) + 2M\epsilon \\
&\leq 3M\epsilon,
\end{aligned}$$

and the uniform convergence is immediate from the uniform Cauchy criterion.

A third consequence can be considered as a uniform version of the result that absolute convergence implies convergence. In practice it tends to be fairly easy to apply, but it applies only in the simplest situations.

Proposition 1.20 (Weierstrass M test). Let S be a nonempty set, and let $\{f_n\}$ be a sequence of real-valued functions on S such that $|f_n(x)| \le M_n$ for all x in S. Suppose that $\sum_n M_n < +\infty$. Then $\sum_{n=1}^\infty f_n(x)$ converges uniformly for x in S.

PROOF. If $n \ge m \ge N$, then $\big|\sum_{k=m+1}^n f_k(x)\big| \le \sum_{k=m}^n |f_k(x)| \le \sum_{k=m}^n M_k$, and the right side tends to 0 uniformly in x as N tends to infinity. Therefore the result follows from the uniform Cauchy criterion.

EXAMPLES.

(1) The series
$$\sum_{n=1}^\infty \frac{1}{n^2}\,x^n$$
converges uniformly for $-1 \le x \le 1$ by the Weierstrass M test with $M_n = 1/n^2$.

(2) The series
$$\sum_{n=1}^\infty (-1)^n\,\frac{x^2+n}{n^2}$$
converges uniformly for $-1 \le x \le 1$, but the M test does not apply. To see that the M test does not apply, we use the smallest possible M_n, which is $M_n = \sup_x \big|(-1)^n \frac{x^2+n}{n^2}\big| = \frac{n+1}{n^2}$. The series $\sum \frac{n+1}{n^2}$ diverges, and hence the M test cannot apply for any choice of the numbers M_n. To see the uniform convergence of the given series, we observe that the terms strictly alternate in sign. Also,
$$\frac{x^2+n}{n^2} \ge \frac{x^2+(n+1)}{(n+1)^2} \quad \text{because} \quad \frac{x^2}{n^2} \ge \frac{x^2}{(n+1)^2} \quad \text{and} \quad \frac{1}{n} \ge \frac{1}{n+1}.$$
Finally
$$\frac{x^2+n}{n^2} \le \frac{n+1}{n^2} \to 0$$
uniformly for $-1 \le x \le 1$. Hence the series converges uniformly by the uniform Leibniz test (Corollary 1.18).

Having developed some tools for proving uniform convergence, let us apply the notion of uniform convergence to interchanges of limits involving functions of a real variable. For a point of reference, recall the diagrams of interchanges of limits at the beginning of the section. We take the column index to be n and think of the row index as a variable t, which is tending to x. We make assumptions that correspond to (i) and (ii) in Proposition 1.16, namely that $\{f_n(t)\}$ converges uniformly in t as n tends to infinity, say to $f(t)$, and that $f_n(t)$ converges to some limit $f_n(x)$ as t tends to x. With $f_n(x)$ defined as this limit, f_n is continuous at x. In other words, the assumptions are that the sequence $\{f_n\}$ is uniformly convergent to f and each f_n is continuous.

Theorem 1.21. If $\{f_n\}$ is a sequence of real-valued functions on $[a, b]$ that are continuous at x and if $\{f_n\}$ converges to f uniformly, then f is continuous at x.

REMARKS. This is really a consequence of Proposition 1.16 except that one of the indices, namely t, is regarded as continuous and not discrete. Actually, there is a subtle simplification here, by comparison with Proposition 1.16, in that $\{f_n(x)\}$ at the limiting parameter x is being assumed to tend to $f(x)$. This corresponds to assuming (b) in the proposition, as well as (i) and (ii). Consequently the proof of the theorem will be considerably simpler than the proof of Proposition 1.16. In fact, the proof will be our first example of a 3ϵ proof. In many applications of Theorem 1.21, the given sequence $\{f_n\}$ is continuous at every x, and then the conclusion is that f is continuous at every x.

PROOF. We write

$$|f(t) - f(x)| \le |f(t) - f_n(t)| + |f_n(t) - f_n(x)| + |f_n(x) - f(x)|.$$

Given $\epsilon > 0$, choose N large enough so that $|f_n(t) - f(t)| < \epsilon$ for all t whenever $n \ge N$. With such an n fixed, choose some δ of continuity for the function f_n, the point x, and the number ϵ. Each term above is then $< \epsilon$, and hence $|f(t) - f(x)| < 3\epsilon$. Since ϵ is arbitrary, f is continuous at x.

Theorem 1.21 in effect uses only conclusion (c) of Proposition 1.16, which concerns the equality of the two iterated limits. Conclusion (d) gives a stronger result, namely that the double limit exists and equals each iterated limit. The strengthened version of Theorem 1.21 is as follows.

Theorem 1.21′. If $\{f_n\}$ is a sequence of real-valued functions on $[a, b]$ that are continuous at x and if $\{f_n\}$ converges to f uniformly, then for each $\epsilon > 0$, there exist an integer N and a number $\delta > 0$ such that

$$|f_n(t) - f(x)| < \epsilon$$

whenever $n \ge N$ and $|t - x| < \delta$.

PROOF. If $\epsilon > 0$ is given, choose N such that $|f_n(t) - f(t)| < \epsilon/2$ for all t whenever $n \ge N$, and choose δ in the conclusion of Theorem 1.21 such that $|t - x| < \delta$ implies $|f(t) - f(x)| < \epsilon/2$. Then

$$|f_n(t) - f(x)| \le |f_n(t) - f(t)| + |f(t) - f(x)| < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon$$

whenever $n \ge N$ and $|t - x| < \delta$. Theorem 1.21′ follows.

In interpreting our diagrams of interchanges of limits to get at the statement of Theorem 1.21, we took the column index to be n and thought of the row index as a variable t, which was tending to x. It is instructive to see what happens when the roles of n and t are reversed, i.e., when the row index is n and the column index is the variable t, which is tending to x. Again we have $f_n(t)$ converging to $f(t)$ and $\lim_{t\to x} f_n(t) = f_n(x)$, but the uniformity is different. This time we want the uniformity to be in n as t tends to x. This means that the δ of continuity that corresponds to ϵ can be taken independent of n. This is the notion of "equicontinuity," and there is a classical theorem about it. The theorem is actually stronger than Proposition 1.16 suggests, since the theorem assumes less than that $f_n(t)$ converges to $f(t)$ for all t.

Let $\mathcal{F} = \{f_\alpha \mid \alpha \in A\}$ be a set of real-valued functions on a bounded interval $[a, b]$. We say that $\mathcal{F}$ is **equicontinuous** at $x \in [a, b]$ if for each $\epsilon > 0$, there is some $\delta > 0$ such that $|t - x| < \delta$ implies $|f(t) - f(x)| < \epsilon$ for all $f \in \mathcal{F}$. The set $\mathcal{F}$ of functions is **pointwise bounded** if for each $t \in [a, b]$, there exists a number M_t such that $|f(t)| \le M_t$ for all $f \in \mathcal{F}$. The set is **uniformly equicontinuous** on $[a, b]$ if it is equicontinuous at each point x and if the δ can be taken independent of x. The set is **uniformly bounded** on $[a, b]$ if it is pointwise bounded at each $t \in [a, b]$ and the bound M_t can be taken independent of t.

Theorem 1.22 (Ascoli's Theorem). If $\{f_n\}$ is a sequence of real-valued functions on a closed bounded interval $[a, b]$ that is equicontinuous at each point of $[a, b]$ and pointwise bounded on $[a, b]$, then

(a) $\{f_n\}$ is uniformly equicontinuous and uniformly bounded on $[a, b]$,
(b) $\{f_n\}$ has a uniformly convergent subsequence.

PROOF. Since each f_n is continuous at each point, we know from Theorems 1.10 and 1.11 that each f_n is uniformly continuous and bounded. The proof of (a) amounts to an argument that the estimates in those theorems can be arranged to apply simultaneously for all n.

First consider the question of uniform boundedness. Choose, by Theorem 1.11, some x_n in $[a, b]$ with $|f_n(x_n)|$ equal to $K_n = \sup_{x\in[a,b]} |f_n(x)|$. Then choose a subsequence on which the numbers K_n tend to $\sup_n K_n$ in $\mathbb{R}^*$. There will be no loss of generality in assuming that this subsequence is our whole sequence. Apply the Bolzano–Weierstrass Theorem to find a convergent subsequence $\{x_{n_k}\}$ of $\{x_n\}$, say with limit x_0. By pointwise boundedness, find M_{x_0} with $|f_n(x_0)| \le M_{x_0}$ for all n. Then choose some δ of equicontinuity at x_0 for $\epsilon = 1$. As soon as k is large enough so that $|x_{n_k} - x_0| < \delta$, we have

$$K_{n_k} = |f_{n_k}(x_{n_k})| \le |f_{n_k}(x_{n_k}) - f_{n_k}(x_0)| + |f_{n_k}(x_0)| < 1 + M_{x_0}.$$

Thus $1 + M_{x_0}$ is a uniform bound for the functions f_n.

The proof of uniform equicontinuity proceeds in the same spirit but takes longer to write out. Fix $\epsilon > 0$. The uniform continuity of f_n for each n means that it makes sense to define

$$\delta_n(\epsilon) = \min\left\{1,\ \sup\left\{\delta' > 0 \;\middle|\; \begin{array}{l} |f(x) - f(y)| < \epsilon \text{ whenever } |x - y| < \delta' \\ \text{and } x \text{ and } y \text{ are in the domain of } f \end{array}\right\}\right\}.$$

If $|x - y| < \delta_n(\epsilon)$, then $|f_n(x) - f_n(y)| < \epsilon$. Put $\delta(\epsilon) = \inf_n \delta_n(\epsilon)$. Let us see that it is enough to prove that $\delta(\epsilon) > 0$: If x and y are in $[a, b]$ with $|x-y| < \delta(\epsilon)$, then $|x - y| < \delta(\epsilon) \le \delta_n(\epsilon)$. Hence $|f_n(x) - f_n(y)| < \epsilon$ as required.

Thus we are to prove that $\delta(\epsilon) > 0$. If $\delta(\epsilon) = 0$, then we first choose an increasing sequence $\{n_k\}$ of positive integers such that $\delta_{n_k}(\epsilon) < \frac{1}{k}$, and we next choose x_k and y_k in $[a, b]$ with $|x_k - y_k| < \delta_{n_k}(\epsilon)$ and $|f_k(x_k) - f_k(y_k)| \ge \epsilon$. Applying the Bolzano–Weierstrass Theorem, we obtain a subsequence $\{x_{k_l}\}$ of $\{x_k\}$ such that $\{x_{k_l}\}$ converges, say to x_0. Then

$$\limsup_l |y_{k_l} - x_0| \le \limsup_l |y_{k_l} - x_{k_l}| + \limsup_l |x_{k_l} - x_0| = 0 + 0 = 0,$$

so that $\{y_{k_l}\}$ converges to x_0. Now choose, by equicontinuity at x_0, a number $\delta' > 0$ such that $|f_n(x) - f_n(x_0)| < \frac{\epsilon}{2}$ for all n whenever $|x - x_0| < \delta'$. The convergence of $\{x_{k_l}\}$ and $\{y_{k_l}\}$ to x_0 implies that for large enough l, we have $|x_{k_l} - x_0| < \delta'/2$ and $|y_{k_l} - x_0| < \delta'/2$. Therefore $|f_{k_l}(x_{k_l}) - f_{k_l}(x_0)| < \frac{\epsilon}{2}$ and $|f_{k_l}(y_{k_l}) - f_{k_l}(x_0)| < \frac{\epsilon}{2}$, from which we conclude that $|f_{k_l}(x_{k_l}) - f_{k_l}(y_{k_l})| < \epsilon$. But we saw that $|f_k(x_k) - f_k(y_k)| \ge \epsilon$ for all k, and thus we have arrived at a contradiction. This proves the uniform equicontinuity and completes the proof of (a).

To prove (b), we first construct a subsequence of $\{f_n\}$ that is convergent at every rational point in $[a, b]$. We enumerate the rationals, say as $x_1, x_2, \ldots$. By the Bolzano–Weierstrass Theorem and the pointwise boundedness, we can find a subsequence of $\{f_n\}$ that is convergent at x_1, a subsequence of the result that is convergent at x_2, a subsequence of the result that is convergent at x_3, and so on. The trouble with this process is that each term of our original sequence may disappear at some stage, and then we are left with no terms that address all the rationals. The trick is to form the subsequence $\{f_{n_k}\}$ of the given $\{f_n\}$ whose k^{th} term is the k^{th} term of the k^{th} subsequence we constructed. Then the k^{th}, $(k+1)^{\text{st}}$, $(k+2)^{\text{nd}}, \ldots$ terms of $\{f_{n_k}\}$ all lie in our k^{th} constructed subsequence, and hence $\{f_{n_k}\}$ converges at the first k points $x_1, \ldots, x_k$. Since k is arbitrary, $\{f_{n_k}\}$ converges at every rational point.

Let us prove that $\{f_{n_k}\}$ is uniformly Cauchy. Let $\epsilon > 0$ be given, let δ be some corresponding number exhibiting equicontinuity, and choose finitely many rationals $r_1, \ldots, r_l$ in $[a, b]$ such that any member of $[a, b]$ is within δ of at least one of these rationals. Then choose N such that $|f_n(r_j) - f_m(r_j)| < \epsilon$ for

$1 \le j \le l$ whenever n and m are $\ge N$. If x is in $[a, b]$, let $r(x)$ be an r_j with $|x - r(x)| < \delta$. Whenever n and m are $\ge N$, we then have

$$\begin{aligned}
&|f_n(x) - f_m(x)| \\
&\qquad \le |f_n(x) - f_n(r(x))| + |f_n(r(x)) - f_m(r(x))| + |f_m(r(x)) - f_m(x)| \\
&\qquad < \epsilon + \epsilon + \epsilon = 3\epsilon.
\end{aligned}$$

Hence $\{f_{n_k}\}$ is uniformly Cauchy, and (b) follows from Proposition 1.17.

REMARK. The construction of the subsequence for which countably many convergence conditions were all satisfied is an important one and is often referred to as a **diagonal process** or as the **Cantor diagonal process**.

EXAMPLE. Let K and M be positive constants, and let $\mathcal{F}$ be the set of continuous real-valued functions f on $[a, b]$ such that $|f(t)| \le K$ for $a \le t \le b$ and such that the derivative $f'(t)$ exists for $a < t < b$ and satisfies $|f'(t)| \le M$ there. This set of functions is certainly uniformly bounded by K, and we show that it is also uniformly equicontinuous. To see the latter, we use the Mean Value Theorem. If x is in the closed interval $[a, b]$ and t is in the open interval (a, b), then there exists ξ depending on t and x such that

$$|f(t) - f(x)| = |f'(\xi)||t - x| \le M|t - x|.$$

From this inequality it follows that the number δ of uniform equicontinuity for ϵ and $\mathcal{F}$ can be taken to be ϵ/M. The hypotheses of Ascoli's Theorem are satisfied, and it follows that any sequence of functions in $\mathcal{F}$ has a uniformly convergent subsequence. The estimate of δ is independent of the uniform bound K, yet Ascoli's Theorem breaks down if there is no bound at all; for example, the sequence of constant functions with $f_n(x) = n$ is uniformly equicontinuous but has no convergent subsequence.

We turn now to the problem of interchange of derivative and limit. The two indices again will be an integer n that is tending to infinity and a parameter t that is tending to x. Proposition 1.16 takes away all the surprise in the statement of the theorem, and it tells us the steps to follow in a proof. What the proposition suggests is that the general entry in our interchange diagram should be whatever quantity we want to take an iterated limit of in either order. Thus we expect not a theorem about a general entry $f_n(t)$, but instead a theorem about a general entry $\dfrac{f_n(t) - f_n(x)}{t - x}$. The limit on n gives us $\dfrac{f(t) - f(x)}{t - x}$ for a limiting function f, and then the limit as $t \to x$ gives us $f'(x)$. In the other order the limit as $t \to x$ gives us $f_n'(x)$, and then we are to consider the limit on n. If Proposition 1.16 is to be a guide, we are to assume that the convergence in one variable is uniform in the other. The proposition also suggests that if we have existence of each row limit and each column limit, then uniform convergence when one variable occurs first is equivalent to uniform convergence when the other variable occurs first. Thus we should assume whichever is easier to verify.

Theorem 1.23. Suppose that $\{f_n\}$ is a sequence of real-valued functions continuous for $a \le t \le b$ and differentiable for $a < t < b$ such that $\{f_n'\}$ converges uniformly for $a < t < b$ and $\{f_n(x_0)\}$ converges in $\mathbb{R}$ for some x_0 with $a \le x_0 \le b$. Then $\{f_n\}$ converges uniformly for $a \le t \le b$ to a function f, and $f'(x) = \lim_n f_n'(x)$ for $a < x < b$, with the derivative and the limit existing.

REMARKS. The convergence of $\{f(x_0)\}$ cannot be dropped completely as a hypothesis because $f_n(t) = n$ would otherwise provide a counterexample. In practice, $\{f_n\}$ will be known in advance to be uniformly convergent. However, uniform convergence of $\{f_n\}$ is not enough by itself, as was shown by the example $f_n(x) = \dfrac{\sin nx}{\sqrt{n}}$ in Section 2.

PROOF. The first step is to apply the Mean Value Theorem to $f_n - f_m$, estimate $f_n' - f_m'$, and use the convergence of $\{f_n(x_0)\}$ to obtain the existence of the limit function f. The Mean Value Theorem produces some ξ strictly between t and x_0 such that

$$f_n(t) - f_m(t) = \big(f_n(x_0) - f_m(x_0)\big) + (t - x_0)\big(f_n'(\xi) - f_m'(\xi)\big).$$

Our hypotheses allow us to conclude that $\{f_n(t)\}$ is uniformly Cauchy, and thus $\{f_n\}$ converges uniformly to a limit function f by Proposition 1.17.

The second step is to apply the Mean Value Theorem again to $f_n - f_m$, this time to see that

$$\varphi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}$$

converges uniformly in t (for $t \ne x$) as n tends to infinity, the limit being $\varphi(t) = \dfrac{f(t) - f(x)}{t - x}$. In fact, the Mean Value Theorem produces some ξ strictly between t and x such that

$$\varphi_n(t) - \varphi_m(t) = \frac{[f_n(t) - f_m(t)] - [f_n(x) - f_m(x)]}{t - x} = f_n'(\xi) - f_m'(\xi),$$

and the right side tends to 0 uniformly as n and m tend to infinity. Therefore $\{\varphi_n(t)\}$ is uniformly Cauchy for $t \ne x$, and Proposition 1.17 shows that it is uniformly convergent.

The third step is to extend the definition of φ to x by $\varphi_n(x) = f_n'(x)$ and then to see that φ_n is continuous at x and Theorem 1.21 applies. In fact, the definition of $\varphi_n(t)$ is as the difference quotient for the derivative of f_n at x, and thus $\varphi_n(t) \to f_n'(x) = \varphi_n(x)$. Hence φ_n is continuous at x. We saw in the second step that $\varphi_n(t)$ is uniformly convergent for $t \neq x$, and we are given that $\varphi_n(x) = f_n'(x)$ is convergent. Therefore $\varphi_n(t)$ is uniformly convergent for all t with

$$\lim \varphi_n(t) = \begin{cases} \dfrac{f(t) - f(x)}{t - x} & \text{for } t \neq x, \\ \lim f_n'(x) & \text{for } t = x. \end{cases}$$

Theorem 1.21 says that the limiting function $\lim \varphi_n(t)$ is continuous at x. Thus

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x} = \lim_n f_n'(x).$$

In other words, f is differentiable at x and $f'(x) = \lim_n f_n'(x)$.

4. Riemann Integral

This section contains a careful but limited development of the Riemann integral in one variable. The reader is assumed to have a familiarity with Riemann sums at the level of a calculus course. The objective in this section is to prove that bounded functions with only finitely many discontinuities are Riemann integrable, to address the interchange-of-limits problem that arises with a sequence of functions and an integration, to prove the Fundamental Theorem of Calculus in the case of continuous integrand, to prove a change-of-variables formula, and to relate Riemann integrals to general Riemann sums. The Riemann integral in several variables will be treated in Chapter III, and some of the theorems to be proved in the several-variable case at that time will be results that have not been proved here in the one-variable case. In Chapters VI and VII, in the context of the Lebesgue integral, we shall prove a much more sweeping version of the Fundamental Theorem of Calculus.

First we give the relevant definitions. We work with a function $f : [a, b] \to \mathbb{R}$ with $a \leq b$ in $\mathbb{R}$, and we always assume that f is bounded. A **partition** P of $[a, b]$ is a subdivision of the interval $[a, b]$ into subintervals, and we write such a partition as

$$a = x_0 \leq x_1 \leq \cdots \leq x_n = b.$$

The points x_j will be called the **subdivision points** of the partition, and we may abbreviate the partition as $P = \{x_i\}_{i=0}^n$. In order to permit integration over an interval of zero length, we allow partitions in which two consecutive x_j's are equal; the **multiplicity** of x_j is the number of times that x_j occurs in the partition. For the above partition, let

$$\Delta x_i = x_i - x_{i-1}, \qquad \mu(P) = \textbf{mesh of } P = \max_i \Delta x_i,$$
$$M_i = \sup_{x_{i-1} \le x \le x_i} f(x), \qquad m_i = \inf_{x_{i-1} \le x \le x_i} f(x).$$

Put

$$U(P, f) = \sum_{i=1}^n M_i \Delta x_i = \text{upper Riemann sum for } P,$$
$$L(P, f) = \sum_{i=1}^n m_i \Delta x_i = \text{lower Riemann sum for } P,$$
$$\overline{\int_a^b} f\,dx = \inf_P U(P, f) = \text{upper Riemann integral of } f,$$
$$\underline{\int_a^b} f\,dx = \sup_P L(P, f) = \text{lower Riemann integral of } f.$$

We say that f is **Riemann integrable** on $[a, b]$ if $\overline{\int_a^b} f\,dx = \underline{\int_a^b} f\,dx$, and in this case we write $\int_a^b f\,dx$ for the common value of these two numbers. We write $\mathcal{R}[a, b]$ for the set of Riemann integrable functions on $[a, b]$.

If $f \ge 0$, an upper Riemann sum for f may be visualized in the traditional way as the sum of the areas of rectangles with bases $[x_{i-1}, x_i]$ and with heights just sufficient to rise above the graph of f on the interval $[x_{i-1}, x_i]$, and a lower sum may be visualized similarly, using rectangles as large as possible so that they lie below the graph.

EXAMPLES.

(1) Suppose $f(x) = c$ for $a \le x \le b$. No matter what partition P is used, we have $M_i = c$ and $m_i = c$. Therefore $U(P, f) = L(P, f) = c(b - a)$, $\overline{\int_a^b} f\,dx = \underline{\int_a^b} f\,dx = c(b - a)$, and f is Riemann integrable on $[a, b]$ with $\int_a^b f\,dx = c(b - a)$.

(2) Let $[a, b]$ be arbitrary with $a < b$, and let f be 1 on the rationals and 0 on the irrationals. This f is discontinuous at every point of $[a, b]$. No matter what partition is used, we have $M_i = 1$ and $m_i = 0$ whenever $\Delta x_i > 0$. Therefore $U(P, f) = b - a$ and $L(P, f) = 0$. Hence $\overline{\int_a^b} f\,dx = b - a$ and $\underline{\int_a^b} f\,dx = 0$, and f is not Riemann integrable.

Let us work toward a proof that continuous functions are Riemann integrable. We shall use some elementary properties of upper and lower Riemann sums along with Theorem 1.10, which says that a continuous function on $[a, b]$ is uniformly continuous.

Lemma 1.24. Suppose that $f : [a, b] \to \mathbb{R}$ has $m \le f(x) \le M$ for all x in $[a, b]$. Then

$$m(b-a) \le L(P, f) \le U(P, f) \le M(b-a),$$

$$m(b-a) \le \underline{\int_a^b} f\,dx \le M(b-a),$$

$$m(b-a) \le \overline{\int_a^b} f\,dx \le M(b-a).$$

PROOF. The first conclusion follows from the computation

$$\begin{aligned} m(b-a) = \sum_{i=1}^n m\Delta x_i \le L(P, f) = \sum_{i=1}^n m_i\,\Delta x_i \\ \le \sum_{i=1}^n M_i\,\Delta x_i = U(P, f) \le \sum_{i=1}^n M\Delta x_i = M(b-a). \end{aligned}$$

If we concentrate on the first, third, and last members of the above inequalities and take the supremum on P, then we obtain the second conclusion. Similarly if we concentrate on the first, sixth, and last members of the above inequalities and take the infimum on P, then we obtain the third conclusion.

A **refinement** of the partition P is a partition P^* containing all the subdivision points of P, with at least their same multiplicities. If P_1 and P_2 are two partitions, then P_1 and P_2 have at least one common refinement: one such common refinement is obtained by taking the union of the subdivision points from each and repeating each such point with the maximum of the multiplicities with which it occurs in P_1 and P_2. We use this notion in order to prove a second lemma.

Lemma 1.25. Let $f : [a, b] \to \mathbb{R}$ satisfy $m \le f(x) \le M$ for all x in $[a, b]$. Then

(a) $L(P, f) \le L(P^*, f)$ and $U(P^*, f) \le U(P, f)$ whenever P is a partition of $[a, b]$ and P^* is a refinement,
(b) $L(P_1, f) \le U(P_2, f)$ whenever P_1 and P_2 are partitions of $[a, b]$,
(c) $\underline{\int_a^b} f\,dx \le \overline{\int_a^b} f\,dx$,

(d) $\overline{\int_a^b} f\,dx - \underline{\int_a^b} f\,dx \le (M-m)(b-a)$,
(e) the function f is Riemann integrable on $[a,b]$ if and only if for each $\epsilon > 0$, there exists a partition P with $U(P,f) - L(P,f) < \epsilon$.

PROOF. In (a), it is enough to handle the case in which P^* is obtained from P by including one additional point, say x^* between x_{i-1} and x_i. The only possible difference between $L(P,f)$ and $L(P^*,f)$ comes from $[x_{i-1},x_i]$, and there we have

$$\begin{aligned}\inf_{x\in[x_{i-1},x_i]} f(x)\,(x_i - x_{i-1}) &= \inf_{x\in[x_{i-1},x_i]} f(x)\,(x_i - x^*) + \inf_{x\in[x_{i-1},x_i]} f(x)\,(x^* - x_{i-1}) \\ &\le \inf_{x\in[x_{i-1},x*]} f(x)\,(x_i - x^*) + \inf_{x\in[x^*,x_i]} f(x)\,(x^* - x_{i-1}).\end{aligned}$$

Hence $L(P,f) \le L(P^*,f)$, and similarly $U(P^*,f) \le U(P,f)$. This proves (a).

Let P^* be a common refinement of P_1 and P_2. Combining (a) with Lemma 1.24 gives

$$L(P_1,f) \le L(P^*,f) \le U(P^*,f) \le U(P_2,f).$$

This proves (b). Conclusion (c) follows by taking the supremum on P_1 and then the infimum on P_2, and conclusion (d) follows by subtracting the second conclusion of Lemma 1.24 from the third.

For (e), we have

$$L(P_1,f) \le \underline{\int_a^b} f\,dx \le \overline{\int_a^b} f\,dx \le U(P_2,f)$$

for any partitions P_1 and P_2 of $[a,b]$. Riemann integrability means that the center two members of this inequality are equal. If they are not equal, then there certainly can exist no P with $U(P,f) - L(P,f) < \epsilon$ if $\epsilon = \overline{\int_a^b} f\,dx - \underline{\int_a^b} f\,dx$. On the other hand, equality of the center two members, together with the definitions of the lower and upper Riemann integrals, means that for each $\epsilon > 0$, we can choose P_1 and P_2 with $U(P_2,f) - L(P_1,f) < \epsilon$. Letting P be a common refinement of P_1 and P_2 and applying (a), we see that $U(P,f) - L(P,f) < \epsilon$. This proves (e).

Theorem 1.26. If $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$, then f is Riemann integrable on $[a,b]$.

PROOF. From Theorem 1.10 we know that f is uniformly continuous on $[a,b]$. Given $\epsilon > 0$, we can therefore choose some number $\delta > 0$ corresponding to f and ϵ on $[a,b]$. Let $P = \{x_i\}_{i=0}^n$ be a partition on $[a,b]$ of mesh $\mu(P) < \delta$. On any subinterval $[x_{i-1}, x_i]$ corresponding to P, we have $m_i = f(\xi_i)$ and $M_i = f(\eta_i)$ for some ξ_i and η_i in $[x_{i-1}, x_i]$, by Theorem 1.11. Since $|\eta_i - \xi_i| \le |x_i - x_{i-1}| = \Delta x_i \le \mu(P) < \delta$, we obtain $M_i - m_i = f(\eta_i) - f(\xi_i) < \epsilon$. Therefore

$$U(P, f) - L(P, f) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i \le \epsilon \sum_{i=1}^{n} \Delta x_i = \epsilon(b-a),$$

and the theorem follows from Lemma 1.25e.

We shall improve upon Theorem 1.26 by allowing finitely many points of discontinuity, but we need to do some additional work beforehand.

Lemma 1.27. If f is bounded on $[a, b]$ and $a \le c \le b$, then $\overline{\int_a^b} f\,dx = \overline{\int_a^c} f\,dx + \overline{\int_c^b} f\,dx$, and similarly for $\underline{\int_a^b}$. Consequently f is in $\mathcal{R}[a, b]$ if and only if f is in both $\mathcal{R}[a, c]$ and $\mathcal{R}[c, b]$, and in this case,

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx.$$

REMARKS. After one is done developing the Riemann integral and its properties, it is customary to adopt the convention that $\int_b^a f\,dx = -\int_a^b f\,dx$ when $b < a$. One of the places that this convention is particularly helpful is in applying the displayed formula of Lemma 1.27: the formula is then valid for all real a, b, c without the assumption that a, b, c are ordered in a particular way.

PROOF. If P_1 and P_2 are partitions of $[a, c]$ and $[c, b]$, respectively, let P be their "union," which is obtained by using all the subdivision points $\ne c$ of each partition, together with c itself. The multiplicity of c in P is to be the larger of the numbers of times c occurs in P_1 and P_2. This P is a partition of $[a, b]$. Then

$$\overline{\int_a^b} f\,dx \le U(P, f) = U(P_1, f) + U(P_2, f).$$

Taking the infimum over P_1 and then the infimum over P_2, we obtain

$$\overline{\int_a^b} f\,dx \le \overline{\int_a^c} f\,dx + \overline{\int_c^b} f\,dx.$$

For the reverse inequality, let $\epsilon > 0$ be given, and choose a partition P of $[a, b]$ with $U(P, f) - \overline{\int_a^b} f\,dx < \epsilon$. Let P^* be the refinement of P obtained by adjoining c to P if c is not a subdivision point of P or by using P itself if c is a subdivision point of P. Lemma 1.25a gives $U(P^*, f) - \overline{\int_a^b} f\,dx < \epsilon$. Because c is a subdivision point of P^*, the subdivision points $\le c$ give us a partition P_1 of $[a, c]$ and the subdivision points $\ge c$ give us a partition P_2 of $[c, b]$. Moreover, P^* is the union of P_1 and P_2. Then we have

$$\overline{\int_a^b} f\,dx + \epsilon \ge U(P^*, f) = U(P_1, f) + U(P_2, f) \ge \overline{\int_a^c} f\,dx + \overline{\int_c^b} f\,dx.$$

Since ϵ is arbitrary, the lemma follows.

Lemma 1.28. Suppose that $f : [a, b] \to \mathbb{R}$ is bounded on $[a, b]$ and that $a \le c \le b$. If for each $\delta > 0$, f is Riemann integrable on each closed subinterval of $[a, b] - \{x \mid |x - c| \ge \delta\}$, then f is Riemann integrable on $[a, b]$.

PROOF. We give the argument when $a < c < b$, the cases $c = a$ and $c = b$ being handled similarly. Since f is by assumption bounded, find m and M with $m \le f(x) \le M$ for all $x \in [a, b]$. Choose $\delta > 0$ small enough so that $a < c - \delta < c < c + \delta < b$. To simplify the notation, let us drop "$f\,dx$" from all integrals. Since f is by assumption Riemann integrable on $[a, c - \delta]$ and $[c + \delta, b]$, Lemma 1.27 gives

$$\begin{aligned}
\overline{\int_a^b} = \overline{\int_a^{c-\delta}} + \overline{\int_{c-\delta}^{c+\delta}} + \overline{\int_{c+\delta}^{b}} &= \underline{\int_a^{c-\delta}} + \overline{\int_{c-\delta}^{c+\delta}} + \underline{\int_{c+\delta}^{b}} \\
&\le \underline{\int_a^{c-\delta}} + \Big(\underline{\int_{c-\delta}^{c+\delta}} + 2\delta(M - m)\Big) + \underline{\int_{c+\delta}^{b}} \\
&= \underline{\int_a^{b}} + 2\delta(M - m).
\end{aligned}$$

Since δ is arbitrary, $\overline{\int_a^b} = \underline{\int_a^b}$. The lemma follows.

Proposition 1.29. If $f : [a, b] \to \mathbb{R}$ is bounded on $[a, b]$ and is continuous at all but finitely many points of $[a, b]$, then f is Riemann integrable on $[a, b]$.

REMARK. There is no assumption that f has only jump discontinuities. For example, the proposition applies if $[a, b] = [0, 1]$ and f is the function with $f(x) = \sin \frac{1}{x}$ for $x \ne 0$ and $f(0) = 0$.

PROOF. By Lemma 1.27 and induction, it is enough to handle the case that f is discontinuous at exactly one point, say c. Since f is bounded and is continuous at all points but c, Theorem 1.26 shows that the hypotheses of Lemma 1.28 are satisfied. Therefore Lemma 1.28 shows that f is Riemann integrable on $[a, b]$.

We shall now work toward a theorem about interchanging limits and integrals. The preliminary step is to obtain some simple properties of Riemann integrals.

Proposition 1.30. If f, f_1, and f_2 are Riemann integrable on $[a, b]$, then

(a) $f_1 + f_2$ is in $\mathcal{R}[a, b]$ and $\int_a^b (f_1 + f_2)\,dx = \int_a^b f_1\,dx + \int_a^b f_2\,dx$,
(b) cf is in $\mathcal{R}[a, b]$ and $\int_a^b cf\,dx = c\int_a^b f\,dx$ for any real number c,
(c) $f_1 \le f_2$ on $[a, b]$ implies $\int_a^b f_1\,dx \le \int_a^b f_2\,dx$,
(d) $m \le f \le M$ on $[a, b]$ and $\varphi : [m, M] \to \mathbb{R}$ continuous imply that $\varphi \circ f$ is in $\mathcal{R}[a, b]$,
(e) $|f|$ is in $\mathcal{R}[a, b]$, and $\left|\int_a^b f\,dx\right| \le \int_a^b |f|\,dx$,
(f) f^2 and $f_1 f_2$ are in $\mathcal{R}[a, b]$,
(g) $\sqrt{f}$ is in $\mathcal{R}[a, b]$ if $f \ge 0$ on $[a, b]$,
(h) the function g with $g(x) = f(-x)$ is in $\mathcal{R}[-b, -a]$ and satisfies $\int_{-b}^{-a} g\,dx = \int_a^b f\,dx$.

REMARK. The proof of (c) will show, even without the assumption of Riemann integrability, that $\overline{\int_a^b} f_1\,dx \le \overline{\int_a^b} f_2\,dx$ and $\underline{\int_a^b} f_1\,dx \le \underline{\int_a^b} f_2\,dx$. We shall make use of this stronger conclusion later in this section.

PROOF. For (a), write $f = f_1 + f_2$, and let P be a partition. From

$$\inf_{x\in[x_{i-1},x_i]} f_1(x) + \inf_{x\in[x_{i-1},x_i]} f_2(x) \le \inf_{x\in[x_{i-1},x_i]} (f_1 + f_2)(x) = \inf_{x\in[x_{i-1},x_i]} f(x)$$

and a similar inequality with the supremum, we obtain

$$L(P, f_1) + L(P, f_2) \le L(P, f) \le U(P, f) \le U(P, f_1) + U(P, f_2). \qquad (*)$$

Let $\epsilon > 0$ be given. By Lemma 1.25e, choose P_1 and P_2 with

$$U(P_1, f_1) - L(P_1, f_1) < \epsilon \quad \text{and} \quad U(P_2, f_2) - L(P_2, f_2) < \epsilon.$$

If P is a common refinement of P_1 and P_2, then Lemma 1.25a gives

$$U(P, f_1) - L(P, f_1) < \epsilon \quad \text{and} \quad U(P, f_2) - L(P, f_2) < \epsilon.$$

Hence

$$U(P, f_1) \le \int_a^b f_1\,dx + \epsilon \le L(P, f_1) + 2\epsilon,$$
$$U(P, f_2) \le \int_a^b f_2\,dx + \epsilon \le L(P, f_2) + 2\epsilon,$$

and $(*)$ yields $U(P, f) - L(P, f) \le 4\epsilon$. Since ϵ is arbitrary, Lemma 1.25e shows that f is in $\mathcal{R}[a, b]$. From the inequalities for $U(P, f_1)$ and $U(P, f_2)$, combined with the last inequality in $(*)$, we see that

$$\int_a^b f\,dx \le U(P, f) \le U(P, f_1) + U(P, f_2) \le \int_a^b f_1\,dx + \int_a^b f_2\,dx + 2\epsilon,$$

while the first inequality in $(*)$ shows that

$$\begin{aligned}\int_a^b f_1\,dx + \int_a^b f_2\,dx + 2\epsilon &\le L(P, f_1) + L(P, f_2) + 4\epsilon \\ &\le L(P, f) + 4\epsilon \le \int_a^b f\,dx + 4\epsilon.\end{aligned}$$

Since ϵ is arbitrary, we obtain $\int_a^b (f_1 + f_2)\,dx = \int_a^b f_1\,dx + \int_a^b f_2\,dx$. This proves (a).

For (b), consider any subinterval $[x_{i-1}, x_i]$ of a partition, and let m_i and M_i be the infimum and supremum of f on this subinterval. Also, let m_i' and M_i' be the infimum and supremum of cf on this subinterval. If $c \ge 0$, then $M_i' = cM_i$ and $m_i' = cm_i$, so that $(M_i' - m_i')\Delta x_i = c(M_i - m_i)\Delta x_i$. If $c \le 0$, then $M_i' = -cm_i$ and $m_i' = -cM_i$, so that we still have $(M_i' - m_i')\Delta x_i = c(M_i - m_i)\Delta x_i$. Summing on i, we obtain $U(P, cf) - L(P, cf) = c(U(P, f) - L(P, f))$, and (b) follows from Lemma 1.25e.

For (c), we have $\overline{\int_a^b} f_1\,dx \le U(P, f_1) \le U(P, f_2)$ for all P. Taking the infimum on P in the inequality of the first and third members gives $\overline{\int_a^b} f_1\,dx \le \overline{\int_a^b} f_2\,dx$. (Similarly $\underline{\int_a^b} f_1\,dx \le \underline{\int_a^b} f_2\,dx$, but this is not needed under the hypothesis that f_1 and f_2 are Riemann integrable.)

For (d), let $K = \sup_{t\in[m,M]} |\varphi(t)|$. Let $\epsilon > 0$ be given, and choose by Theorem 1.10 some δ of uniform continuity for φ and ϵ. Without loss of generality, we may assume that $\delta \le \epsilon$. By Lemma 1.25e, choose a partition $P = \{x_i\}_{i=0}^n$ of $[a, b]$ such that $U(P, f) - L(P, f) < \delta^2$. On any subinterval $[x_{i-1}, x_i]$ of P, let m_i and M_i be the infimum and supremum of f, and let m_i' and M_i' be the infimum and supremum of $\varphi \circ f$. Divide the set of integers $\{1, \dots, n\}$ into two subsets—the subset A of integers i with $M_i - m_i < \delta$ and the subset B of integers i with $M_i - m_i \ge \delta$. If i is in A, then the definition of δ makes $M_i' - m_i' \le \epsilon$. If i is in B, then the best we can say is that $M_i' - m_i' \le 2K$. However, on B we do have $M_i - m_i \ge \delta$, and thus

$$\delta \sum_{i\in B} \Delta x_i \le \sum_{i\in B} (M_i - m_i)\Delta x_i \le \sum_{i=1}^n (M_i - m_i)\Delta x_i = U(P,f) - L(P,f) < \delta^2.$$

Thus $\sum_{i\in B} \Delta x_i < \delta$ and

$$\begin{aligned} U(P, \varphi\circ f) - L(P, \varphi\circ f) &= \sum_{i\in A} (M_i' - m_i')\Delta x_i + \sum_{i\in B} (M_i' - m_i')\Delta x_i \\ &\le \epsilon \sum_{i\in A} \Delta x_i + 2K \sum_{i\in B} \Delta x_i \\ &\le \epsilon(b-a) + 2K\delta \le \epsilon(b-a) + 2K\epsilon. \end{aligned}$$

Since ϵ is arbitrary, the Riemann integrability of $\varphi\circ f$ follows from Lemma 1.25e.

For (e), the first conclusion follows from (d) with $\varphi(t) = |t|$. For the asserted inequality we have $f \le |f|$ and $-f \le |f|$, so that (c) and (b) give $\int_a^b f\,dx \le \int_a^b |f|\,dx$ and $-\int_a^b f\,dx \le \int_a^b |f|\,dx$. Combining these inequalities, we obtain $\big|\int_a^b f\,dx\big| \le \int_a^b |f|\,dx$.

For (f), the first conclusion follows from (d) with $\varphi(t) = t^2$. For the Riemann integrability of $f_1 f_2$, we use the formula $f_1 f_2 = \frac{1}{2}\big((f_1+f_2)^2 - f_1^2 - f_2^2\big)$ and the earlier parts of the proposition.

Conclusion (g) follows from (d) with $\varphi(t) = \sqrt{t}$.

For (h), each partition P of $[a, b]$ yields a natural partition P' of $[-b, -a]$ by using the negatives of the partition points. When P and P' are matched in this way, $U(P, f) = U(P', g)$ and $L(P, f) = L(P', g)$. It is immediate that $f \in \mathcal{R}[a, b]$ implies $g \in \mathcal{R}[-b, -a]$ and that $\int_{-b}^{-a} g\,dx = \int_a^b f\,dx$. This completes the proof.

The next topic is the problem of interchange of integral and limit.

EXAMPLE. On the interval $[0, 1]$, define $f_n(x)$ to be n for $0 < x < 1/n$ and to be 0 otherwise. Proposition 1.29 shows that f_n is Riemann integrable, and Lemma 1.27 allows us to see that $\int_0^1 f_n\,dx = 1$ for all n. On the other hand, $\lim_n f_n(x) = 0$ for all $x \in [0, 1]$. Since $\int_0^1 0\,dx = 0$, we have $\int_0^1 f_n\,dx = 1 \ne 0 = \lim_n \int f_n\,dx$. Thus the interchange is not justified without some additional hypothesis.

Theorem 1.31. If $\{f_n\}$ is a sequence of Riemann integrable functions on $[a, b]$ and if $\{f_n\}$ converges uniformly to f on $[a, b]$, then f is Riemann integrable on $[a, b]$, and $\lim_n \int_a^b f_n\,dx = \int_a^b f\,dx$.

REMARKS. Proposition 1.16 suggests considering a "matrix" whose entries are the quantities for which we are computing iterated limits, and these quantities are $U(P, f_n)$ here. (Alternatively, we could use $L(P, f_n)$.) The hypothesis of uniformity in the statement of Theorem 1.31, however, concerns f_n, not $U(P, f_n)$. In fact, the tidy hypothesis on f_n in the statement of the theorem implies a less intuitive hypothesis on $U(P, f_n)$ that has not been considered. The proof conceals these details.

PROOF. Using the uniform Cauchy criterion with $\epsilon = 1$, we see that there exists N such that $|f_n(x)| \le M_N + 1$ for all x whenever $n \ge N$. It follows from the boundedness of $f_1, \dots, f_{N-1}$ that the $|f_n|$ are uniformly bounded, say by M. Then also $|f(x)| \le M$ for all x. Put $\varepsilon_n = \sup_x |f_n(x) - f(x)|$, so that $f_n - \varepsilon_n \le f \le f_n + \varepsilon_n$. Proposition 1.30c and the remark with the proposition, combined with Lemma 1.25c, then yield

$$\int_a^b (f_n - \varepsilon_n)\,dx \le \underline{\int_a^b} f\,dx \le \overline{\int_a^b} f\,dx \le \int_a^b (f_n + \varepsilon_n)\,dx.$$

Hence $\overline{\int_a^b} f\,dx - \underline{\int_a^b} f\,dx \le 2\varepsilon_n(b-a)$ for all n. The uniform convergence of $\{f_n\}$ to f forces ε_n to tend to 0, and thus f is in $\mathcal{R}[a,b]$. The displayed equation, in light of the Riemann integrability of f, shows that

$$\Big| \int_a^b f\,dx - \int_a^b f_n\,dx \Big| \le 2\varepsilon_n(b-a).$$

The right side tends to 0, and therefore $\lim_n \int_a^b f_n\,dx = \int_a^b f\,dx$.

EXAMPLE. Let $f : [0,1] \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} 1/q & \text{if } x \text{ is the rational } p/q \text{ in lowest terms} \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

This function is discontinuous at every rational and is continuous at every irrational. Its Riemann integrability is not settled by Proposition 1.29. Define

$$f_n(x) = \begin{cases} 1/q & \text{if } x \text{ is the rational } p/q \text{ in lowest terms, } q \le n \\ 0 & \text{if } x \text{ is the rational } p/q \text{ in lowest terms, } q > n \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

Proposition 1.29 shows that f_n is Riemann integrable, and Lemma 1.27 shows that $\int_0^1 f_n\,dx = 0$. Since $|f_n(x) - f(x)| \le 1/n$ for all x, $\{f_n\}$ converges uniformly to f. By Theorem 1.31, f is Riemann integrable and $\int_0^1 f\,dx = 0$.

Theorem 1.32 (Fundamental Theorem of Calculus). If $f : [a,b] \to \mathbb{R}$ is continuous, then

(a) the function $G(x) = \int_a^x f\,dt$ is differentiable for $a < x < b$ with derivative $f(x)$, and it is continuous at a and b with $G(a) = 0$,

(b) any continuous function F on $[a,b]$ that is differentiable for $a < x < b$ with derivative $f(x)$ has $\int_a^b f\,dt = F(b) - F(a)$.

REMARK. The derivative of $G(x)$ on (a, b), namely $f(x)$, has the finite limits $f(a)$ and $f(b)$ at the endpoints of the interval, since f has been assumed to be continuous on $[a, b]$. Thus, in the sense of the last paragraph of Section A2 of the appendix, $G(x)$ has the continuous derivative $f(x)$ on the *closed* interval $[a, b]$.

PROOF OF (a). Riemann integrability of f is known from Theorem 1.26. For $h > 0$ small enough to make $x + h < b$, Lemma 1.27 and Proposition 1.30 give

$$\begin{aligned} \frac{G(x+h) - G(x)}{h} - f(x) &= \frac{\int_a^{x+h} f\,dt - \int_a^x f\,dt}{h} - f(x) \\ &= \frac{1}{h} \int_x^{x+h} f\,dt - f(x) \\ &= \frac{1}{h} \int_x^{x+h} [f(t) - f(x)]\,dt \end{aligned}$$

and hence $$\left| \frac{G(x+h) - G(x)}{h} - f(x) \right| \le \frac{1}{h} \int_x^{x+h} |f(t) - f(x)|\,dt.$$

If $\epsilon > 0$ is given, choose the δ of continuity for f and ϵ at x. Then $0 < h \le \delta$ implies that the right side is $\le \epsilon$. For negative h, we instead take $h > 0$ and consider

$$\frac{G(x-h) - G(x)}{-h} - f(x) = \frac{1}{h} \int_{x-h}^{x} f\,dt - f(x) = \frac{1}{h} \int_{x-h}^{x} [f(t) - f(x)]\,dt.$$

Then

$$\left| \frac{G(x-h) - G(x)}{-h} - f(x) \right| \le \frac{1}{|h|} \int_{x-h}^{x} |f(t) - f(x)|\,dt \le \epsilon,$$

as required.

PROOF OF (b). The functions F and G are two continuous functions on $[a, b]$ with equal derivative on (a, b). A corollary of the Mean Value Theorem stated in Section A2 of the appendix implies $G = F + c$ for some constant c. Then

$$\int_a^b f\,dt = G(b) - 0 = G(b) - G(a) = F(b) + c - F(a) - c = F(b) - F(a).$$

Corollary 1.33 (integration by parts). Let f and g be real-valued functions defined and having a continuous derivative on $[a, b]$. Then

$$\int_a^b f(x) g'(x)\,dx = \Big[f(x) g(x) \Big]_a^b - \int_a^b f'(x) g(x)\,dx.$$

REMARK. The notion of a continuous derivative at the endpoints of an interval is discussed in the last paragraph of Section A2 of the appendix.

PROOF. We start from the product rule for differentiation, namely

$$\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + f'(x)g(x),$$

and we apply $\int_a^b$ to both sides. Taking Theorem 1.32 into account, we obtain the desired formula.

Theorem 1.34 (change-of-variables formula). Let f be Riemann integrable on $[a, b]$, let φ be a continuous strictly increasing function from an interval $[A, B]$ onto $[a, b]$, suppose that the inverse function $\varphi^{-1} : [a, b] \to [A, B]$ is continuous, and suppose finally that φ is differentiable on (A, B) with uniformly continuous derivative φ'. Then the product $(f \circ \varphi)\varphi'$ is Riemann integrable on $[A, B]$, and

$$\int_a^b f(x)\,dx = \int_A^B f(\varphi(y))\varphi'(y)\,dy.$$

REMARKS. The uniform continuity of φ' forces φ' to be bounded. If φ' were also assumed positive on (A, B), then the continuity of φ^{-1} on (a, b) would be automatic as a consequence of the proposition in Section A3 of the appendix. The result in the appendix is not quite good enough for current purposes, and thus we have assumed the continuity of φ^{-1} on $[a, b]$. It will be seen in Section II.7 that the continuity of φ^{-1} on $[a, b]$ is automatic in the statement of Theorem 1.34 and need not be assumed.

PROOF IF $f \geq 0$. Given $\epsilon > 0$, choose some η of uniform continuity for φ' and ϵ, and then choose, by Theorem 1.10, some δ of uniform continuity for φ^{-1} and η. Next choose a partition $P = \{x_i\}_{i=0}^n$ on $[a, b]$ such that $U(P, f) - L(P, f) < \epsilon$. Possibly by passing to a refinement of P, we may assume that $\mu(P) < \delta$. Let Q be the partition $\{y_i\}_{i=0}^n$ of $[A, B]$ with $y_i = \varphi^{-1}(x_i)$. Then $\mu(Q) < \eta$.

The Mean Value Theorem gives $\Delta x_i = (\Delta y_i)\varphi'(\xi_i)$ for some ξ between y_{i-1} and y_i. On $[A, B]$, φ' is bounded; let m_i^* and M_i^* be the infimum and supremum of φ' on $[y_{i-1}, y_i]$, so that $m_i^* \leq \varphi'(\xi) \leq M_i^*$ and $m_i^*\Delta y_i \leq \Delta x_i \leq M_i^*\Delta y_i$. Since $\mu(Q) < \eta$, we have $M_i^* - m_i^* \leq \epsilon$. Then we have

$$\sum M_i m_i^* \Delta y_i \leq \sum M_i \Delta x_i = U(P, f) = \sum M_i \varphi'(\xi_i)\Delta y_i \leq M_i M_i^* \Delta y_i.$$

Whenever F and G are ≥ 0 on a common domain and x is in that domain, $(\inf G)F(x) \leq G(x)F(x)$; taking the supremum of both sides gives the inequality $(\inf G)(\sup F) \leq \sup(FG)$. Also, $\sup(FG) \leq \sup(F)\sup(G)$. Applying these inequalities with $G = \varphi'$ and $F = f \circ \varphi$ yields

$$\sum M_i m_i^* \Delta y_i \leq U(Q, (f \circ \varphi)\varphi') \leq M_i M_i^* \Delta y_i.$$

Subtraction of the right-hand inequality of the first display and the left-hand inequality of the second display shows that

$$U(P, f) - U(Q, (f \circ \varphi)\varphi') \leq \sum M_i (M_i^* - m_i^*) \Delta y_i,$$

while subtraction of the right-hand inequality of the second display and the left-hand inequality of the first display gives

$$U(Q, (f \circ \varphi)\varphi') - U(P, f) \leq \sum M_i (M_i^* - m_i^*) \Delta y_i.$$

Therefore

$$|U(P, f) - U(Q, (f \circ \varphi)\varphi')| \leq \sum M_i (M_i^* - m_i^*) \Delta y_i \leq \epsilon M(B - A).$$

Similarly
$$|L(P, f) - L(Q, (f \circ \varphi)\varphi')| \leq \epsilon M(B - A),$$

and hence

$$\begin{aligned} &|U(Q, (f \circ \varphi)\varphi') - L(Q, (f \circ \varphi)\varphi')| \\ &\qquad \leq |U(Q, (f \circ \varphi)\varphi') - U(P, f)| + |U(P, f) - L(P, f)| \\ &\qquad\quad + |L(P, f) - L(Q, (f \circ \varphi)\varphi')| \\ &\qquad \leq 2\epsilon M(B - A) + \epsilon. \end{aligned}$$

Since ϵ is arbitrary, Lemma 1.25e shows that $(f \circ \varphi)\varphi'$ is in $\mathcal{R}[A, B]$. Our inequalities imply that

$$\begin{aligned} \Big|\textstyle\int_a^b f\,dx - U(P, f)\Big| &\leq \epsilon, \\ |U(P, f) - U(Q, (f \circ \varphi)\varphi')| &\leq \epsilon M(B - A), \end{aligned}$$

and
$$\Big|U(Q, (f \circ \varphi)\varphi') - \textstyle\int_A^B (f \circ \varphi)\varphi'\,dy\Big| \leq 2\epsilon M(B - A).$$

Addition shows that $\Big|\int_a^b f\,dx - \int_A^B (f \circ \varphi)\varphi'\,dy\Big| \leq \epsilon + 3\epsilon M(B - A)$. Since ϵ is arbitrary, $\int_a^b f\,dx = \int_A^B (f \circ \varphi)\varphi'\,dy$.

Proof for general f. The special case just proved shows that the result holds for $f + c$ for a suitable positive constant c, as well as for the constant function c. Subtracting the results for $f + c$ and c gives the result for f, and the proof is complete.

If f is Riemann integrable on $[a, b]$, then $U(P, f)$ and $L(P, f)$ tend to $\int_a^b f\,dx$ as P gets finer by insertion of points. This conclusion tells us nothing about fine-looking partitions like those that are equally spaced with many subdivisions. The next theorem says that the approximating sums tend to $\int_a^b f\,dx$ just under the assumption that $\mu(P)$ tends to 0.

Relative to our standard partition $P = \{x_i\}_{i=0}^n$, let t_i for $1 \le i \le n$ satisfy $x_{i-1} \le t_i \le x_i$, and define

$$S(P, \{t_i\}, f) = \sum_{i=1}^n f(t_i)\Delta x_i.$$

This is called a **Riemann sum** of f.

Theorem 1.35. If f is Riemann integrable on $[a, b]$, then

$$\lim_{\mu(P)\to 0} S(P, \{t_i\}, f) = \int_a^b f\,dx.$$

Conversely if f is bounded on $[a, b]$ and if there exists a real number r such that for any $\epsilon > 0$, there exists some $\delta > 0$ for which $|S(P, \{t_i\}, f) - r| < \epsilon$ whenever $\mu(P) < \delta$, then f is Riemann integrable on $[a, b]$.

PROOF. For the direct part the function f is assumed bounded; suppose $|f(x)| \le M$ on $[a, b]$. Let $\epsilon > 0$ be given. Choose a partition P^* of $[a, b]$ with $U(P^*, f) \le \int_a^b f\,dx + \epsilon$. Say P^* is a partition into k intervals. Put $\delta_1 = \frac{\epsilon}{Mk}$, and suppose that P is any partition of $[a, b]$ with $\mu(P) \le \delta_1$. In the sum giving $U(P, f)$, we divide the terms into two types—those from a subinterval of P that does not lie within one subinterval of P^* and those from a subinterval of P that does lie within one subinterval of P^*.

Each subinterval of P of the first kind has at least one point of P^* strictly inside it, and the number of such subintervals is therefore $\le k - 1$. Hence the sum of the corresponding terms of $U(P, f)$ is

$$\le (k-1)M\mu(P) \le \frac{(k-1)M\epsilon}{Mk} \le \epsilon.$$

The sum of the terms of the second kind is $\le U(P^*, f)$. Thus

$$U(P, f) \le \epsilon + U(P^*, f) \le \int_a^b f\,dx + 2\epsilon.$$

Similarly we can produce δ_2 such that $\mu(P) \le \delta_2$ implies

$$L(P, f) \ge \int_a^b f\,dx - 2\epsilon.$$

If $\delta = \min\{\delta_1, \delta_2\}$ and $\mu(P) \le \delta$, then

$$\int_a^b f\,dx - 2\epsilon \le L(P, f) \le S(P, f) \le U(P, f) \le \int_a^b f\,dx + 2\epsilon,$$

and hence $\left|S(P, f) - \int_a^b f\,dx\right| \le 2\epsilon$.

For the converse let $\epsilon > 0$ be given, and choose some δ as in the statement of the theorem. Next choose a partition $P = \{x_i\}_{i=0}^n$ with $\left|U(P, f) - \overline{\int_a^b} f\,dx\right| < \epsilon$ and $\left|\underline{\int_a^b} f\,dx - L(P, f)\right| < \epsilon$; possibly by passing to a refinement of P, we may assume without loss of generality that $\mu(P) < \delta$. Choosing $\{t_i\}_{i=1}^n$ suitably for the partition P, we can make $|U(P, f) - S(P, \{t_i\}, f)| < \epsilon$. For a possibly different choice of the set of intermediate points, say $\{t_i'\}$, we can make $|S(P, \{t_i'\}, f) - L(P, f)| < \epsilon$. Then

$$\begin{aligned}\left|\overline{\int_a^b} f\,dx - \underline{\int_a^b} f\,dx\right| &\le \left|U(P, f) - \overline{\int_a^b} f\,dx\right| + |U(P, f) - S(P, \{t_i\}, f)| \\ &\quad + |S(P, \{t_i\}, f) - r| + |r - S(P, \{t_i'\}, f)| \\ &\quad + |S(P, \{t_i'\}, f) - L(P, f)| + \left|L(P, f) - \underline{\int_a^b} f\,dx\right| \\ &< 6\epsilon.\end{aligned}$$

Since ϵ is arbitrary, the Riemann integrability of f follows from Lemma 1.25e.

With integration in hand, one could at this point give rigorous definitions of the logarithm and exponential functions $\log x$ and $\exp x$, as well as rigorous but inconvenient definitions of the trigonometric functions $\sin x$, $\cos x$, and $\tan x$. For each of these functions we would obtain a formula for the derivative and other information. We shall not pursue this approach, but we pause to mention the idea. We put $\log x = \int_1^x t^{-1}\,dt$ for $0 < x < +\infty$ and see that log carries $(0, +\infty)$ one-one onto $(-\infty, +\infty)$. The function $\log x$ has derivative $1/x$ and satisfies the functional equation $\log(xy) = \log x + \log y$. The proposition in Section A3 of the appendix shows that the inverse function exp exists, carries $(-\infty, +\infty)$ one-one onto $(0, +\infty)$, is differentiable, and has derivative $\exp x$. The functional equation of log translates into the functional equation $\exp(a + b) = \exp a \exp b$ for exp, and we readily derive as a consequence that $\exp x = e^x$, where $e = \exp 1$. For the trigonometric functions, the starting points with this approach are the definitions $\arctan x = \int_0^x (1 + t^2)^{-1}\,dt$, $\arcsin x = \int_0^x (1 - t^2)^{-1/2}\,dt$, and $\pi = 4 \arctan 1$.

Instead of using this approach, we shall use power series to define these functions and to obtain their expected properties. We do so in Section 7.

5. Complex-Valued Functions

Complex numbers are taken as known, and their notation and basic properties are reviewed in Section A4 of the appendix. The point of the present section is to extend some of the results for real-valued functions in earlier sections so that they apply also to complex-valued functions.

The **distance** between two members z and w of $\mathbb{C}$ is defined by $d(z, w) = |z - w|$. This has the properties

(i) $d(z_1, z_2) \geq 0$ with equality if and only if $z_1 = z_2$,
(ii) $d(z_1, z_2) = d(z_2, z_1)$,
(iii) $d(z_1, z_2) \leq d(z_1, z_3) + d(z_3, z_2)$.

The first two are immediate from the definition, and the third follows from the triangle inequality of Section A4 of the appendix with $z = z_1 - z_3$ and $w = z_3 - z_2$. For this reason, (iii) is called the **triangle inequality** also.

Convergence of a sequence $\{z_n\}$ in $\mathbb{C}$ to z has two possible interpretations: either $\{\operatorname{Re} z_n\}$ converges to $\operatorname{Re} z$ and $\{\operatorname{Im} z_n\}$ converges to $\operatorname{Im} z$, or $d(z_n, z)$ converges to 0 in $\mathbb{R}$. These interpretations come to the same thing because

$$\max\{\operatorname{Re} w, \operatorname{Im} w\} \leq |w| \leq \sqrt{2}\max\{\operatorname{Re} w, \operatorname{Im} w\}.$$

Then it follows that uniform convergence of a sequence of complex-valued functions has two equivalent meanings, so does continuity of a complex-valued function at a point or everywhere, and so does differentiation of a complex-valued function. We readily check that all the results of Section 3, starting with Proposition 1.16 and ending with Theorem 1.23, extend to be valid for complex-valued functions as well as real-valued functions.

The one point that requires special note in connection with Section 3 is the Mean Value Theorem. This theorem is valid for real-valued functions but not for complex-valued functions. It is possible to give an example now if we again allow ourselves to use the exponential and trigonometric functions before we get to Section 7, where the tools will be available for rigorous definitions. The example is $f(x) = e^{ix}$ for $x \in [0, 2\pi]$. This function has $f(0) = f(2\pi) = 1$, but the derivative $f'(x) = ie^{ix}$ is never 0.

The Mean Value Theorem was used in the proof of Theorem 1.23, but the failure of the Mean Value Theorem for complex-valued functions causes us no problem when we seek to extend Theorem 1.23 to complex-valued functions. The reason is that once Theorem 1.23 has been proved for real-valued functions, one simply puts together conclusions about the real and imaginary parts.

Next we examine how the results of Section 4 may be extended to complex-valued functions. Upper and lower Riemann sums, of course, make no sense for a complex-valued function. It is possible to make sense out of general Riemann sums as in Theorem 1.35, but we shall not base a definition on this approach.

Instead, we simply define definite integrals of a function $f : \mathbb{R} \to \mathbb{C}$ in terms of real and imaginary parts. Define the real and imaginary parts $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ by $f(x) = u(x) + iv(x)$, and let $\int_a^b f\,dx = \int_a^b u\,dx + i\int_a^b v\,dx$. We can then redefine the set $\mathcal{R}[a, b]$ of **Riemann integrable** functions on $[a, b]$ to consist of bounded complex-valued functions on $[a, b]$ whose real and imaginary parts are each Riemann integrable.

Most properties of definite integrals go over to the case of complex-valued functions by inspection; there are two properties that deserve some discussion:

(i) If f is in $\mathcal{R}[a, b]$ and c is complex, then cf is in $\mathcal{R}[a, b]$ and $\int_a^b cf\,dx = c\int_a^b f\,dx$.

(ii) If f is in $\mathcal{R}[a, b]$, then $|f|$ is in $\mathcal{R}[a, b]$ and $\left|\int_a^b f\,dx\right| \le \int_a^b |f|\,dx$.

To see (i), write $f = u + iv$ and $c = p + iq$. Then $cf = (r + is)(u + iv) = (ru - sv) + i(rv + su)$. The functions $ru - sv$ and $rv + su$ are Riemann integrable on $[a, b]$, and hence so is cf. Then

$$\begin{aligned}\int_a^b cf\,dx &= \int_a^b (ru - sv)\,dx + i\int_a^b (rv + su)\,dx\\ &= r\int_a^b u\,dx - s\int_a^b v\,dx + ir\int_a^b v\,dx + is\int_a^b u\,dx\\ &= r\int_a^b (u + iv)\,dx + is\int_a^b (u + iv)\,dx = c\int_a^b f\,dx.\end{aligned}$$

To see (ii), let f be in $\mathcal{R}[a, b]$. Proposition 1.30 shows successively that $(\operatorname{Re} f)^2$ and $(\operatorname{Im} f)^2$ are in $\mathcal{R}[a, b]$, that $(\operatorname{Re} f)^2 + (\operatorname{Im} f)^2 = |f|^2$ is in $\mathcal{R}[a, b]$, and that $\sqrt{|f|^2} = |f|$ is in $\mathcal{R}[a, b]$. For the inequality with $\left|\int_a^b f\,dx\right|$, choose $c \in \mathbb{C}$ with $|c| = 1$ such that $c\int_a^b f\,dx$ is real and nonnegative, i.e., equals $\left|\int_a^b f\,dx\right|$. Using (i), we obtain (ii) from

$$\begin{aligned}\left|\int_a^b f\,dx\right| &= c\int_a^b f\,dx = \int_a^b cf\,dx = \int_a^b \operatorname{Re}(cf)\,dx\\ &\le \int_a^b |cf|\,dx = \int_a^b |f|\,dx.\end{aligned}$$

Finally we observe that Theorem 1.35 extends to complex-valued functions f. The definition of **Riemann sum** is unchanged, namely $S(P, \{t_i\}, f) = \sum_{i=1}^n f(t_i)\Delta x_i$, and the statement of Theorem 1.35 is unchanged except that the number r is now allowed to be complex. The direct part of the extended theorem follows by applying Theorem 1.35 to the real and imaginary parts of f separately. For the converse we use that the inequality $|S(P, \{t_i\}, f) - c| < \epsilon$ with c complex implies $|S(P, \{t_i\}, \operatorname{Re} f) - \operatorname{Re} c| < \epsilon$ and $|S(P, \{t_i\}, \operatorname{Im} f) - \operatorname{Im} c| < \epsilon$. Theorem 1.35 for real-valued functions then shows that $\operatorname{Re} f$ and $\operatorname{Im} f$ are Riemann integrable, and hence so is f.

6. Taylor's Theorem with Integral Remainder

There are several forms to the remainder term in the one-variable Taylor's Theorem for real-valued functions, and the differences already show up in their lowest-order formulations. Let f be given, and, for definiteness, suppose $a < x$. If $o(1)$ denotes a term that tends to 0 as x tends to a, three such lowest-order formulas are

$f(x) = f(a) + o(1)$	if f is merely assumed to be continuous,
$f(x) = f(a) + (x-a)f'(\xi)$	with $a < \xi < x$ if f is continuous on $[a, x]$ and f' exists on (a, x),
$f(x) = f(a) + \int_a^x f'(t)\,dt$	if f and f' are continuous on $[a, x]$.

The first formula follows directly from the definition of continuity, while the second formula restates the Mean Value Theorem and the third formula restates part of the Fundamental Theorem of Calculus. The hypotheses of the three formulas increase in strength, and so do the conclusions. In practice, Taylor's Theorem is most often used with functions having derivatives of all orders, and then the strongest hypothesis is satisfied. Thus we state a general theorem corresponding only to the third formula above. It applies to complex-valued functions as well as real-valued functions.

Theorem 1.36 (Taylor's Theorem). Let n be an integer ≥ 0, let a and x be points of $\mathbb{R}$, and let f be a complex-valued function with $n+1$ continuous derivatives on the closed interval from a to x. Then

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(a, x),$$

where

$$R_n(a, x) = \begin{cases} \dfrac{1}{n!}\displaystyle\int_a^x (x-t)^n f^{(n+1)}(t)\,dt & \text{if } a \leq x, \\ -\dfrac{1}{n!}\displaystyle\int_x^a (x-t)^n f^{(n+1)}(t)\,dt & \text{if } x \leq a. \end{cases}$$

REMARKS. The notion of a continuous derivative at the endpoints of an interval is discussed for real-valued functions in the last paragraph of Section A2 of the appendix and extends immediately to complex-valued functions; iteration of this definition attaches a meaning to continuous higher-order derivatives on a closed interval. Once the convention in the remarks with Lemma 1.27 is adopted, namely that $\int_x^a f\,dt = -\int_a^x f\,dt$ when $x < a$, the formula for the remainder term becomes tidier:

$$R_n(a,x) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt,$$

with no assumption that $a \le x$.

PROOF. We give the argument when $a \le x$, the case $x \le a$ being handled analogously. The proof is by induction on n. For $n = 0$, the formula is immediate from the Fundamental Theorem of Calculus (Theorem 1.32b). Assume that the formula holds for $n-1$. We apply integration by parts (Corollary 1.33) to the remainder term at stage $n-1$, obtaining

$$\int_a^x (x-t)^{n-1} f^{(n)}(t)\,dt = -\frac{1}{n}\Big[(x-t)^n f^{(n)}(t)\Big]_a^x + \frac{1}{n}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt.$$

Substitution gives

$$\begin{aligned} R_{n-1}(a,x) &= \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f^{(n)}(t)\,dt \\ &= -\frac{1}{n!}\Big[(x-t)^n f^{(n)}(t)\Big]_a^x + \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt \\ &= \frac{1}{n!}(x-a)^n f^{(n)}(a) + R_n(a,x), \end{aligned}$$

and the induction is complete.

7. Power Series and Special Functions

A **power series** is an infinite series of the form $\sum_{n=0}^\infty c_n z^n$. Normally in mathematics, if nothing is said to the contrary, the coefficients c_n are assumed to be complex and the variable z is allowed to be complex. However, in the context of real-variable theory, as when forming derivatives of functions defined on intervals, one is interested only in real values of z. In this book the context will generally make clear whether the variable is to be regarded as complex or as real.

One source of power series is the "infinite Taylor series" $\sum_{n=0}^\infty \dfrac{f^{(n)}(0)x^n}{n!}$ of a function f having derivatives of all orders, with the remainder terms discarded. In this case the variable is to be real. If the series is convergent at x, the series has sum $f(x)$ if and only if $\lim_n R_n(0,x) = 0$. Later in this section, we shall see examples where the limit is identically 0 and where it is nowhere 0 for $x \ne 0$.

Theorem 1.37. If a power series $\sum_{n=0}^{\infty} c_n z^n$ is convergent in $\mathbb{C}$ for some complex z_0 with $|z_0| = R$ and if $R' < R$, then $\sum_{n=0}^{\infty} |c_n z^n|$ is uniformly convergent for complex z with $|z| \le R'$, and so is $\sum_{n=0}^{\infty} (n+1)|c_{n+1} z^n|$.

REMARKS. The number

$$R = \sup \left\{ R' \,\middle|\, \textstyle\sum_{n=0}^{\infty} c_n z^n \text{ converges for some } z_0 \text{ with } |z_0| = R' \right\}$$

is called the **radius of convergence** of $\sum_{n=0}^{\infty} c_n z^n$. The theorem says that if $R' < R$, then $\sum_{n=0}^{\infty} |c_n z^n|$ converges uniformly for $|z| \le R'$, and it follows from the uniform Cauchy criterion that $\sum_{n=0}^{\infty} c_n z^n$ converges uniformly for $|z| \le R'$. The definition of R carries with it the implication that if z_0 has $|z_0| > R$, then $\sum_{n=0}^{\infty} c_n z_0^n$ diverges.

PROOF. The theorem is vacuous unless $R > 0$. Since $\sum_{n=0}^{\infty} c_n z_0^n$ is convergent, the terms $c_n z_0^n$ tend to 0. Thus there is some integer N for which $|c_n| R^n \le 1$ when $n \ge N$. Fix $R' < R$. For $|z| \le R'$ and $n \ge N$, we have

$$|c_n z^n| = |c_n z_0^n| \left| \frac{z}{z_0} \right|^n = |c_n| R^n \left| \frac{z}{z_0} \right|^n \le \left(\frac{R'}{R} \right)^n.$$

Since $\sum \left(\frac{R'}{R} \right)^n < +\infty$, the Weierstrass M test shows that $\sum_{n=0}^{\infty} c_n z^n$ converges uniformly for $|z| \le R'$.

For the series $\sum_{n=0}^{\infty} (n+1)|c_{n+1} z^n|$, the inequalities $|z| \le R'$ and $n \ge N$ together imply

$$|(n+1) c_{n+1} z^n| \le (n+1) |c_{n+1}| R^n \left(\frac{R'}{R} \right)^n \le (n+1) R^{-1} \left(\frac{R'}{R} \right)^n.$$

To see that the Weierstrass M test applies here as well, choose r' with $R'/R < r' < 1$ and increase the size of N so that $\frac{n+1}{n} \le \frac{R}{R'} r'$ whenever $n \ge N$. For such n, the ratio test and the inequality

$$\frac{(n+2) R^{-1} \left(\frac{R'}{R} \right)^{n+1}}{(n+1) R^{-1} \left(\frac{R'}{R} \right)^n} = \frac{n+2}{n+1} \frac{R'}{R} \le r'$$

show that $\sum (n+1) R^{-1} \left(\frac{R'}{R} \right)^n$ converges. Thus the Weierstrass M test indeed applies, and the proof is complete.

Corollary 1.38. If $\sum_{n=0}^{\infty} c_n x^n$ converges for $|x| < R$ and the sum of the series for x real is denoted by $f(x)$, then the function f has derivatives of all orders for $|x| < R$. These derivatives are given by term-by-term differentiation of the series for f, and each differentiated series converges for $|x| < R$. Moreover, $c_k = \dfrac{f^{(k)}(0)}{k!}$.

REMARK. When a function has derivatives of all orders, we say that it is **infinitely differentiable**.

PROOF. The corollary is vacuous unless $R > 0$. Let $R' < R$. The given series certainly converges at $x = 0$, and Theorem 1.37 shows that the term-by-term differentiated series converges uniformly for $|x| \leq R'$. Thus Theorem 1.23 gives $f'(x) = \sum_{n=0}^{\infty} (n+1)c_{n+1}x^n$ for $|x| < R'$. Since $R' < R$ is arbitrary, $f'(x) = \sum_{n=0}^{\infty} (n+1)c_{n+1}x^n$ for $|x| < R$.

We can iterate this result to obtain the corresponding conclusion for the higher-order derivatives. Evaluating the derivatives at 0, we obtain $f^{(k)}(0) = c_k k!$, as asserted.

Corollary 1.39. If $\sum_{n=0}^{\infty} c_n x^n$ and $\sum_{n=0}^{\infty} d_n x^n$ both converge for $|x| < R$ and if their sums are equal for x real with $|x| < R$, then $c_n = d_n$ for all n.

PROOF. This result is immediate from the formula for the coefficients in Corollary 1.38.

If $f : \mathbb{R} \to \mathbb{C}$ is infinitely differentiable near $x = a$, we call the infinite series $\sum\limits_{n=0}^{\infty} \dfrac{f^{(n)}(a)}{n!}(x-a)^n$ the (infinite) **Taylor series** of f. We call a series $\sum_{n=0}^{\infty} c_n(x-a)^n$ a **power series** about $x = a$; its behavior at $x = a + t$ is the same as the behavior of the series $\sum_{n=0}^{\infty} c_n x^n$ at $x = t$. In applications, one usually adjusts the function f so that Taylor series expansions are about $x = 0$. Thus we shall concentrate largely on power series expansions about $x = 0$.

Had we chosen at the end of Section 4 to define $\log x$ as $\int_1^x t^{-1}\,dt$ and $\exp x$ as the inverse function of $\log x$, we would have found right away that $\left(\frac{d}{dx}\right)^k \exp x = \exp x$ for all k. Therefore the infinite Taylor series expansion of $\exp x$ about $x = 0$ is $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. This fact does not, however, tell us whether $\exp x$ is the sum of this series. For this purpose we need to examine the remainder. Theorem 1.36 shows that the remainder after the term $x^n/n!$ is

$$R_n(0, x) = \frac{1}{n!}\int_0^x (x-t)^n f^{(n+1)}(t)\,dt = \frac{1}{n!}\int_0^x (x-t)^n e^t\,dt.$$

Between 0 and x, e^t is bounded by some constant $M(x)$ depending on x, and thus $|R_n(0, x)| \leq \frac{M(x)}{n!}\left|\int_0^x (x-t)^n\,dt\right| = \frac{M(x)}{(n+1)!}|x|^{n+1}$. With x fixed, this tends to 0 as n tends to infinity, and thus $\lim_n R_n(0, x) = 0$ for each x. The conclusion is that $\exp x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. In a similar fashion one can obtain power series expansions of $\sin x$ and $\cos x$ if one starts from definitions of the corresponding inverse functions in terms of Riemann integrals.

Instead of using this approach, we shall define $\exp x$, $\sin x$, and $\cos x$ directly as sums of standard power series. An advantage of using series in the definitions is that this approach allows us to define these functions at an arbitrary complex z, not just at a real x. Thus we define

$$\exp z = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \qquad \sin z = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!}, \qquad \cos z = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!}.$$

The ratio test shows immediately that these series all converge for all complex z. Inspection of all these series gives us the identity

$$\exp iz = \cos z + i \sin z.$$

Corollary 1.38 shows that the functions $\exp z$, $\sin z$, and $\cos z$, when considered as functions of a real variable $z = x$, are infinitely differentiable with derivatives given by the expected formulas

$$\frac{d}{dx} \exp x = \exp x, \qquad \frac{d}{dx} \sin x = \cos x, \qquad \frac{d}{dx} \cos x = -\sin x.$$

From these formulas it is immediate that $\frac{d}{dx}\big(\sin^2 x + \cos^2 x\big) = 0$ for all x. Therefore $\sin^2 x + \cos^2 x$ is constant. Putting $x = 0$ shows that the constant is 1. Thus

$$\sin^2 x + \cos^2 x = 1.$$

In order to prove that $\exp x = e^x$, where $e = \exp 1$, and to prove other familiar trigonometric identities, we shall do some calculations with power series that are justified by the following theorem.

Theorem 1.40. If $f(z) = \sum_{n=0}^{\infty} a_n z^n$ and $g(z) = \sum_{n=0}^{\infty} b_n z^n$ for complex z with $|z| < R$, then $f(z)g(z) = \sum_{n=0}^{\infty} c_n z^n$ for $|z| < R$, where

$$c_n = a_n b_0 + a_{n-1} b_1 + \cdots + a_0 b_n.$$

REMARK. In other words, the rule is to multiply the series formally, assuming a kind of infinite distributive law, and reassemble the series by grouping terms with like powers of z. The coefficient c_n of z^n in the product comes from all products $a_k z^k b_l z^l$ for which the total degree is n, i.e., for which $k + l = n$. Thus c_n is as indicated.

PROOF. The theorem is vacuous unless $R > 0$. Fix $R' < R$. For $|z| \le R'$, put $F(z) = \sum_{n=0}^{\infty} |a_n z^n|$ and $G(z) = \sum_{n=0}^{\infty} |b_n z^n|$. These series are uniformly convergent for $|z| \le R'$ by Theorem 1.37, and also $|f(z)| \le F(z)$ and $|g(z)| \le G(z)$. By the uniform convergence of the series for F and G when $|z| \le R'$, there exists $M < +\infty$ such that $F(z) \le M$ and $G(z) \le M$ for $|z| \le R'$. Given

$\epsilon > 0$, choose an integer N' such that $|z| \le R'$ implies $\sum_{n \ge N'} |a_n z^n| < \epsilon$ and $\sum_{n \ge N'} |b_n z^n| < \epsilon$. If $|z| \le R'$ and $N \ge 2N'$, then

$$\Big| f(z)g(z) - \sum_{n=0}^{\infty} c_n z^n \Big| \le \Big| f(z)g(z) - \Big(\sum_{n=0}^{N} a_n z^n\Big)\Big(\sum_{n=0}^{N} b_n z^n\Big)\Big| + \Big|\Big(\sum_{n=0}^{N} a_n z^n\Big)\Big(\sum_{n=0}^{N} b_n z^n\Big) - \sum_{n=0}^{N} c_n z^n\Big|.$$

Call the two terms on the right side T_1 and T_2. Then we have

$$T_1 \le \Big| f(z) - \sum_{n=0}^{N} a_n z^n \Big| |g(z)| + \Big| \sum_{n=0}^{N} a_n z^n \Big| \Big| g(z) - \sum_{n=0}^{N} b_n z^n \Big| \le \epsilon G(z) + \epsilon F(z),$$

and also, with $[N/2]$ denoting the greatest integer in $N/2$,

$$\begin{aligned} T_2 = \Big| \sum_{\substack{k+l>N,\\ k \le N,\ l \le N}} a_k b_l z^{k+l} \Big| &\le \sum_{\substack{k+l>N,\\ k \le N,\ l \le N}} |a_k z^k| |b_l z^l| \\ &\le \sum_{k=0}^{N} \sum_{l=[N/2]}^{N} + \sum_{k=[N/2]}^{N} \sum_{l=0}^{N} \\ &\le \sum_{k=0}^{\infty} \sum_{l=N'}^{\infty} + \sum_{k=N'}^{\infty} \sum_{l=0}^{\infty} \\ &\le \epsilon G(z) + \epsilon F(z). \end{aligned}$$

Since $G(z) \le M$ and $F(z) \le M$ for $|z| \le R'$, the total estimate is that $T_1 + T_2 \le 4\epsilon M$. Since ϵ is arbitrary, we conclude that $\lim_N \sum_{n=0}^{N} c_n z^n = f(z)g(z)$ for $|z| \le R'$. Since R' is an arbitrary number $< R$, we conclude that $\sum_{n=0}^{\infty} c_n z^n = f(z)g(z)$ for $|z| < R$.

Corollary 1.41. For any z and w in $\mathbb{C}$, $\exp(z+w) = \exp z \exp w$. Furthermore, $\exp \bar{z} = \overline{\exp z}$.

PROOF. Theorem 1.40 and the infinite radius of convergence allow us to write

$$\begin{aligned} \exp z \exp w &= \Big(\sum_{r=0}^{\infty} \frac{z^r}{r!}\Big)\Big(\sum_{s=0}^{\infty} \frac{w^s}{s!}\Big) = \sum_{r,s} \frac{z^r w^s}{r!s!} \\ &= \sum_{N=0}^{\infty} \sum_{k=0}^{N} \frac{z^k w^{N-k}}{k!(N-k)!} = \sum_{N=0}^{\infty} \frac{1}{N!} \sum_{k=0}^{N} \binom{N}{k} z^k w^{N-k} \\ &= \sum_{N=0}^{\infty} \frac{(z+w)^N}{N!} = \exp(z+w). \end{aligned}$$

For the second formula, write $z = x + iy$. Then

$$\begin{aligned}\exp \bar{z} &= \exp(x - iy) = \exp x \exp(-iy) = (\exp x)(\cos y - i \sin y) \\ &= \overline{(\exp x)(\cos y + i \sin y)} = \overline{\exp x \exp(iy)} = \overline{\exp(x + iy)} = \overline{\exp z}.\end{aligned}$$

Corollary 1.42. The exponential function $\exp x$, as a function of a real variable, has the following properties:

(a) exp is strictly increasing on $(-\infty, +\infty)$ and is one-one onto $(0, +\infty)$,
(b) $\exp x = e^x$, where $e = \exp 1$,
(c) $\exp x$ has an inverse function, denoted by $\log x$, that is strictly increasing, carries $(0, +\infty)$ one-one onto $(-\infty, +\infty)$, has derivative $1/x$, and satisfies $\log(xy) = \log x + \log y$.

REMARKS. The three facts that $\exp x = e^x$ for x real, $\exp z$ satisfies the functional equation of Corollary 1.41 for z complex, and e^z is previously undefined for z nonreal allow one to define e^z to mean $\exp z$ for all complex z. We follow this convention. In particular, $e^{ix} = \exp(ix) = \cos x + i \sin x$.

PROOF. For $x \geq 0$, we certainly have $\exp x \geq 1$. Also, each term of the series for $\exp x$ is strictly increasing for $x \geq 0$, and hence the same thing is true of the sum of the series. From Corollary 1.41, $\exp(-x) \exp x = \exp 0 = 1$, and thus $\exp x$ is strictly increasing for $x \leq 0$ with $0 < \exp x \leq 1$. Putting these statements together, we see that $\exp x$ is strictly increasing and positive on $(-\infty, +\infty)$. Hence it is one-one. This proves part of (a).

Since $\exp x > 0$, it makes sense to consider rational powers of $\exp x$. Iteration of the identity $\exp(z + w) = \exp z \exp w$ shows that $\big(\exp \frac{px}{q}\big)^q = \exp px = (\exp x)^p$, and application of the q^{th} root function gives $\exp \frac{px}{q} = (\exp x)^{p/q}$. Taking $x = 1$ yields $\exp(p/q) = e^{p/q}$ for all rational p/q. The two functions $\exp x$ and e^x are continuous functions of a real variable that are equal when x is rational, and hence they are equal for all x. This proves (b).

From the first two terms of the series for $\exp 1$, we see that $e > 2$. Therefore $e^n > 2^n > n$ for all positive integers n, and $\exp x$ has arbitrarily large numbers in its image. The Intermediate Value Theorem (Theorem 1.12) then shows that $[0, +\infty)$ is contained in its image. Since $\exp(-x) \exp x = 1$, the interval $(0, 1]$ is contained in the image as well. Thus $\exp x$ carries $(-\infty, +\infty)$ onto $(0, +\infty)$. This proves the remainder of (a).

Consequently $\exp x$ has an inverse function, which is denoted by $\log x$. Since $\exp x$ has the continuous everywhere-positive derivative $\exp x$, the proposition in Section A3 of the appendix applies and shows that $\log x$ is differentiable with derivative $1/\exp(\log x)$. Since exp and log are inverse functions, $\exp(\log x) = x$. Thus the derivative of $\log x$ is $1/x$.

Finally $\exp(\log x + \log y) = \exp(\log x)\exp(\log y) = xy$, since exp and log are inverse functions. Applying log to both sides gives $\log x + \log y = \log(xy)$. This proves (c).

Corollary 1.43. The trigonometric functions $\sin x$ and $\cos x$, as functions of a real variable, satisfy

(a) $\sin(x + y) = \sin x \cos y + \cos x \sin y$,
(b) $\cos(x + y) = \cos x \cos y - \sin x \sin y$.

PROOF. By Corollary 1.41, $\cos(x + y) + i \sin(x + y) = e^{i(x+y)} = e^{ix}e^{iy} = (\cos x + i \sin x)(\cos y + i \sin y)$. Multiplying out the right side and equating real and imaginary parts yields the corollary.

The final step in the foundational work with the trigonometric functions is to define π and to establish the role that it plays with trigonometric functions.

Proposition 1.44. The function $\cos x$, with x real, has a smallest positive x_0 for which $\cos x_0 = 0$. If π is defined by writing $x_0 = \pi/2$, then

(a) $\sin x$ is strictly increasing, hence one-one, from $[0, \frac{\pi}{2}]$ onto $[0, 1]$, and $\cos x$ is strictly decreasing, hence one-one, from $[0, \frac{\pi}{2}]$ onto $[0, 1]$,
(b) $\sin(-x) = -\sin x$ and $\cos(-x) = \cos x$,
(c) $\sin(x + \frac{\pi}{2}) = \cos x$ and $\cos(x + \frac{\pi}{2}) = -\sin x$,
(d) $\sin(x + \pi) = -\sin x$ and $\cos(x + \pi) = -\cos x$,
(e) $\sin(x + 2\pi) = \sin x$ and $\cos(x + 2\pi) = \cos x$.

PROOF. The function $\cos x$ has $\cos 0 = 1$. Arguing by contradiction, suppose that $\cos x$ is nowhere 0 for $x > 0$. By the Intermediate Value Theorem (Theorem 1.12), $\cos x > 0$ for $x \geq 0$. Since $\sin x$ is 0 at 0 and has derivative $\cos x$, $\sin x$ is strictly increasing for $x \geq 0$ and is therefore positive for $x > 0$. Since $\cos x$ has derivative $-\sin x$, $\cos x$ is strictly decreasing for $x \geq 0$. Let us form the function $f(x) = (\cos x - \cos 1) + (\sin 1)(x - 1)$. If there is some $x_1 > 1$ with $f(x_1) > 0$, then the Mean Value Theorem produces some ξ with $1 < \xi < x_1$ such that

$$0 < f(x_1) = f(x_1) - f(1) = (x_1 - 1)f'(\xi) = (x_1 - 1)(-\sin \xi + \sin 1) < 0,$$

and we have a contradiction. Thus $f(x) \leq 0$ for all $x \geq 1$. In other words, $\cos x \leq \cos 1 - (\sin 1)(x - 1)$ for all $x \geq 1$. For x sufficiently large, $\cos 1 - (\sin 1)(x - 1)$ is negative, and we see that $\cos x$ has to be negative for x sufficiently large. The result is a contradiction, and we conclude that $\cos x$ is 0 for some $x > 0$. Let x_0 be the infimum of the nonempty set of positive x's for which $\cos x = 0$. We can find a sequence $\{x_n\}$ with $x_n \to x_0$ and $\cos x_n = 0$ for all n. By continuity

$\cos x_0 = 0$. We know that $x_0 \geq 0$, and we must have $x_0 > 0$, since $\cos 0 = 1$. This proves the existence of x_0.

Since $\sin x$ has derivative $\cos x$, which is positive for $0 \leq x < \pi/2$, $\sin x$ is strictly increasing for $0 \leq x \leq \pi/2$. From $\sin^2 x + \cos^2 x = 1$, we deduce that $\sin(\pi/2) = 1$. By the Intermediate Value Theorem, $\sin x$ is one-one from $[0, \frac{\pi}{2}]$ onto $[0, 1]$. In similar fashion, $\cos x$ is strictly decreasing and one-one from $[0, \frac{\pi}{2}]$ onto $[0, 1]$. This proves (a).

Conclusion (b) is immediate from the series expansions of $\sin x$ and $\cos x$. Conclusion (c) follows from Corollary 1.43 and the facts that $\sin \frac{\pi}{2} = 1$ and $\cos \frac{\pi}{2} = 0$. Conclusion (d) follows by applying (c) twice, and conclusion (e) follows by applying (d) twice.

Corollary 1.45. The function e^{ix}, with x real, has $|e^{ix}| = 1$ for all x, and $x \mapsto e^{ix}$ is one-one from $[0, 2\pi)$ onto the unit circle of $\mathbb{C}$, i.e., the subset of $z \in \mathbb{C}$ with $|z| = 1$.

PROOF. We have $|\cos x + i \sin x|^2 = \cos^2 x + \sin^2 x = 1$ and therefore $|e^{ix}| = 1$. If $e^{ix_1} = e^{ix_2}$ with x_1 and x_2 in $[0, 2\pi)$, then $e^{i(x_1 - x_2)} = 1$ with $t = x_1 - x_2$ in $(-2\pi, 2\pi)$. So $\cos t = 1$ and $\sin t = 0$. From Proposition 1.44 we see that the only possibility for $t \in (-2\pi, 2\pi)$ is $t = 0$. Thus $x_1 - x_2 = 0$, and $x \mapsto e^{ix}$ is one-one.

Now let $x + iy$ have $x^2 + y^2 = 1$. First suppose that $x \geq 0$ and $y \geq 0$. Since $0 \leq y \leq 1$, it follows that there exists $t \in [0, \frac{\pi}{2}]$ with $\sin t = y$. For this t, the numbers x and $\cos t$ are both ≥ 0 and have square equal to $1 - y^2$. Thus $x = \cos t$ and $e^{it} = x + iy$. For a general $x + iy$ with $x^2 + y^2 = 1$, exactly one of the complex numbers $i^n(x + iy)$ with $0 \leq n \leq 3$ has real and imaginary parts ≥ 0. Then $i^n(x + iy) = e^{it}$ for some t. Since $i = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} = e^{i\pi/2}$, we see that $x + iy = e^{it}e^{-in\pi/2} = e^{it - in\pi/2}$. From $e^{i(t \pm 2\pi)} = e^{it}$, we can adjust $it - in\pi/2$ additively by a multiple of $2\pi i$ so that the result it' lies in $i[0, 2\pi)$, and then $e^{it'} = x + iy$, as required.

Corollary 1.46.

(a) The function $\sin x$ carries $(-\frac{\pi}{2}, \frac{\pi}{2})$ onto $(-1, 1)$, has everywhere-positive derivative, and has a differentiable inverse function $\arcsin x$ carrying $(-1, 1)$ one-one onto $(-\frac{\pi}{2}, \frac{\pi}{2})$. The derivative of $\arcsin x$ is $1/\sqrt{1 - x^2}$.

(b) The function $\tan x = (\sin x)/\cos x)$ carries $(-\frac{\pi}{2}, \frac{\pi}{2})$ onto $(-\infty, +\infty)$, has everywhere-positive derivative, and has a differentiable inverse function $\arctan x$ carrying $(-\infty, +\infty)$ one-one onto $(-\frac{\pi}{2}, \frac{\pi}{2})$. The derivative of $\arctan x$ is $1/(1 + x^2)$, and $\int_{-1}^{1} (1 + x^2)^{-1}\, dx = \pi/2$.

PROOF. From Proposition 1.44 we see that $\frac{d}{dx}(\sin x) = \cos x$ and $\frac{d}{dx}(\tan x) = (\cos x)^{-2}$. The first of these is everywhere positive because of (a) and (b) in the proposition, and the second is everywhere positive by inspection. The image of $\sin x$ is $(-1, 1)$ by (a) and (b), and also the image of $\tan x$ is $(-\infty, +\infty)$ by (a) and (b). Application of the proposition in Section A3 of the appendix yields all the conclusions of the corollary except the formula for $\int_{-1}^{1} (1 + x^2)^{-1}\, dx$. This integral is given by $\arctan 1 - \arctan(-1)$ by Theorem 1.32. Since $\tan(\pi/4) = \sin(\pi/4)/\cos(\pi/4)$, (c) in the proposition gives $\tan(\pi/4) = 1$, and hence $\arctan 1 = \pi/4$. In addition, $\tan(-\pi/4) = \frac{\sin(-\pi/4)}{\cos(-\pi/4)} = -\frac{\sin(\pi/4)}{\cos(\pi/4)} = -1$, and hence $\arctan(-1) = -\pi/4$. Therefore $\int_{-1}^{1} (1+x^2)^{-1}\, dx = (\pi/4) - (-\pi/4) = \pi/2$.

A power series, even a Taylor series, may have any radius of convergence in $[0, +\infty]$. Even if the radius of convergence is > 0, the series may not converge to the given function. For example, Problems 20–22 at the end of the chapter ask one to verify that the function

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases}$$

is infinitely differentiable, even at $x = 0$, and has $f^{(n)}(0) = 0$ for all n. Thus its infinite Taylor series is identically 0, and the series evidently converges to $f(x)$ only for $x = 0$.

Because of Corollary 1.38, one is not restricted to a rote use of Taylor's formula in order to compute Taylor series. If we are interested in the Taylor expansion of f about $x = 0$, any power series with a positive radius of convergence that converges to f on some open interval about x has to be the Taylor expansion of f. A simple example is e^{x^2}, whose derivatives at $x = 0$ are a chore to compute. However, $e^u = \sum_{n=0}^{\infty} \frac{u^n}{n!}$ for all u. If we put $u = x^2$, we obtain $e^{x^2} = \sum_{n=0}^{\infty} \frac{x^{2n}}{n!}$ for all x. Therefore this series must be the infinite Taylor series of e^{x^2}. Here is a more complicated example.

EXAMPLE. **Binomial series**. Let p be any complex number, and put $F(x) = (1+x)^p$ for $-1 < x < 1$. We can compute the n^{th} derivative of f by inspection, and we obtain $F^{(n)}(x) = p(p-1)\cdots(p-n+1)(1+x)^{p-n}$. Therefore the infinite Taylor series of F about $x = 0$ is

$$\sum_{n=0}^{\infty} \frac{p(p-1)\cdots(p-n+1)}{n!}\, x^n.$$

This series reduces to a polynomial if p is a nonnegative integer, and the series is genuinely infinite otherwise. The ratio test shows that the series converges for $|x| < 1$; let $f(x)$ be its sum for x real. The remainder term $R_n(0, x)$ is difficult to estimate, and thus the relationship between the sum $f(x)$ and the original function $(1+x)^p$ is not immediately apparent. However, we can use Corollary 1.38 to obtain

$$f'(x) = \sum_{n=1}^{\infty} \frac{np(p-1)\cdots(p-n+1)}{n!}\, x^{n-1} = \sum_{n=0}^{\infty} \frac{p(p-1)\cdots(p-n)}{n!}\, x^n$$

for $|x| < 1$. We compute $(1+x)f'(x)$ by multiplying the first series by x, the second series by 1, and adding. If we write the constant term separately, the result is

$$(1+x)f'(x) = p + \sum_{n=1}^{\infty} \frac{p(p-1)\cdots(p-n+1)[n+(p-n)]}{n!}\, x^n = pf(x).$$

Therefore

$$\begin{aligned}\frac{d}{dx}\big[(1+x)^{-p}f(x)\big] &= -p(1+x)^{-p-1}f(x) + (1+x)^{-p}f'(x)\\ &= (1+x)^{-p-1}[-pf(x) + (1+x)f'(x)] = 0,\end{aligned}$$

and $(1+x)^{-p}f(x)$ has to be constant for $|x| < 1$. From the series whose sum is $f(x)$, we see that $f(0) = 1$, and hence the constant is 1. Thus $f(x) = (1+x)^p$, and we have established the binomial series expansion

$$(1+x)^p = \sum_{n=0}^{\infty} \frac{p(p-1)\cdots(p-n+1)}{n!}\, x^n$$

for $-1 < x < 1$.

8. Summability

Summability refers to an operation on a sequence of complex numbers to make it more likely that the sequence will converge. The subject is of interest particularly with Fourier series, where the ordinary partial sums may not converge even at points where the given function is continuous.

Let $\{s_n\}_{n\geq 0}$ be a sequence in $\mathbb{C}$, and define its sequence of **Cesàro sums**, or arithmetic means, to be given by

$$\sigma_n = \frac{s_0 + s_1 + \cdots + s_n}{n+1}$$

for $n \geq 0$. If $\lim_n \sigma_n = \sigma$ exists in $\mathbb{C}$, we say that $\{s_n\}$ is **Cesàro summable** to the limit σ. For example the sequence with $s_n = (-1)^n$ for $n \geq 0$ is not convergent, but it is Cesàro summable to the limit 0 because σ_n is 0 for all odd n and is $\frac{1}{n+1}$ for all even n.

Theorem 1.47. If a complex sequence $\{s_n\}_{n\geq 0}$ is convergent in $\mathbb{C}$ to the limit s, then $\{s_n\}$ is Cesàro summable to the limit s.

REMARK. The argument is a 2ϵ proof, and two things are affecting σ_n. For k small and fixed, the contribution of s_k to σ_n is $s_k/(n+1)$ and is tending to 0. For k large, any s_k is close to s, and the average of such terms is close to s.

PROOF. Let $\epsilon > 0$ be given, and choose N_1 such that $k \geq N_1$ implies $|s_k - s| < \epsilon$. If $n \geq N_1$, then

$$\sigma_n - s = \frac{(s_0 - s) + \cdots + (s_{N_1} - s)}{n+1} + \frac{(s_{N_1+1} - s) + \cdots + (s_n - s)}{n+1},$$

so that

$$\begin{aligned}|\sigma_n - s| &\leq \frac{|s_0| + \cdots + |s_{N_1}| + (N_1+1)|s|}{n+1} + \frac{n - N_1}{n+1}\epsilon \\ &\leq \frac{|s_0| + \cdots + |s_{N_1}| + (N_1+1)|s|}{n+1} + \epsilon.\end{aligned}$$

The numerator of the first term is fixed, and thus we can choose $N \geq N_1$ large enough so that the first term is $< \epsilon$ whenever $n \geq N$. If $n \geq N$, then we see that $|\sigma_n - s| < 2\epsilon$. Since ϵ is arbitrary, the theorem follows.

Next let $\{a_n\}_{n\geq 0}$ be a complex sequence, and let $\{s_n\}_{n\geq 0}$ be the sequence of partial sums with $s_n = \sum_{k=0}^n a_k$. Form the power series $\sigma_r = \sum_{n=0}^\infty a_n r^n$. We say that the sequence $\{s_n\}$ of partial sums is **Abel summable** to the limit s in $\mathbb{C}$ if $\lim_{r\uparrow 1} \sigma_r = s$, i.e., if for each $\epsilon > 0$, there is some r_0 such that $r_0 \leq r < 1$ implies that $|\sigma_r - s| < \epsilon$. For example, take $a_k = (-1)^k$, so that s_n equals 1 if n is even and equals 0 if n is odd. The sequence $\{s_n\}$ of partial sums is divergent. The r^{th} Abel sum σ_r is given by the geometric series $\sum_{k=0}^\infty (-1)^k r^k$ with sum $1/(1+r)$. Letting r increase to 1, we see that $\{s_n\}$ is Abel summable with limit $\frac{1}{2}$.

Theorem 1.48 (Abel's Theorem). Let $\{a_n\}_{n\geq 0}$ be a complex sequence, and let $\{s_n\}_{n\geq 0}$ be the sequence of partial sums with $s_n = \sum_{k=0}^n a_k$. If $\{s_n\}_{n\geq 0}$ is convergent in $\mathbb{C}$ to the limit s, then $\{s_n\}$ is Abel summable to the limit s.

REMARK. The proof will proceed along the same lines as in the previous case. It is first necessary to express the Abel sums σ_r in terms of the s_k's.

PROOF. Since $\{s_n\}$ converges, $\{s_n\}$ and $\{a_n\}$ are bounded, and thus $\sum_{n=0}^\infty s_n r^n$ and $\sum_{k=0}^\infty a_k r^k$ are absolutely convergent for $0 \leq r < 1$. With $s_{-1} = 0$, write

$$\begin{aligned}\sigma_r &= \sum_{n=0}^\infty a_n r^n = \sum_{n=0}^\infty (s_n - s_{n-1}) r^n = \sum_{n=0}^\infty s_n r^n - \sum_{n=0}^\infty s_n r^{n+1} \\ &= (1-r)\sum_{n=0}^\infty s_n r^n = (1-r)\sum_{k=0}^N r^k s_k + \sum_{k=N+1}^\infty (1-r) r^k s_k.\end{aligned}$$

Let $\epsilon > 0$ be given, and choose N such that $k \geq N$ implies $|s_k - s| < \epsilon$. Then

$$\begin{aligned}
|\sigma_r - s| &\leq (1-r)\sum_{k=0}^{N}(|s_k| + |s|) + \sum_{k=N+1}^{\infty}(1-r)r^k|s_k - s| \\
&\leq (1-r)\sum_{k=0}^{N}(|s_k| + |s|) + \Big((1-r)\sum_{k=N+1}^{\infty} r^k\Big)\epsilon \\
&\leq (1-r)\sum_{k=0}^{N}(|s_k| + |s|) + \epsilon.
\end{aligned}$$

With N fixed, the coefficient of $(1-r)$ in the first term is fixed, and thus we can choose r_0 close enough to 1 so that the first term is $< \epsilon$ whenever $r_0 \leq r < 1$. If $r_0 \leq r < 1$, we see that $|\sigma_r - s| < 2\epsilon$. Since ϵ is arbitrary, the theorem follows.

EXAMPLE. For $|x| < 1$, the geometric series $\sum_{n=0}^{\infty}(-1)^n x^n$ converges and has sum $(1+x)^{-1}$. The Fundamental Theorem of Calculus gives $\log(1+t) = \int_0^x \frac{1}{1+t}\,dt = \int_0^x \sum_{n=0}^{\infty}(-1)^n t^n\,dt$ for $|x| < 1$, and Theorem 1.31 allows us to interchange sum and integral as long as $|x| < 1$. Consequently

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

for $|x| < 1$. The sequence of partial sums on the right converges for $x = 1$ by the Leibniz test, and Theorem 1.48 says that the Abel sums must converge to the same limit. But the Abel sums have limit $\lim_{x\uparrow 1}\log(1+x) = \log 2$, since $\log(1+x)$ is continuous for $x > 0$. Thus Abel's Theorem has given us a rigorous proof of the familiar identity

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{n+1} = \log 2.$$

Theorems 1.47 and 1.48, which say that one kind of convergence always implies another, are called **Abelian theorems**. Converse results, saying that the second kind of convergence implies the first under an additional hypothesis, are called **Tauberian theorems**. These tend to be harder to prove. We give two examples of Tauberian theorems; the first one will be applied immediately to yield an important special case of the main theorem of Section 9; the second one will be used in Chapter VI to prove a deep theorem about pointwise convergence of Fourier series.

Proposition 1.49. Let $\{a_n\}_{n\geq 0}$ be a complex sequence with all terms ≥ 0, and let $\{s_n\}_{n\geq 0}$ be the sequence of partial sums with $s_n = \sum_{k=0}^n a_k$. If $\{s_n\}_{n\geq 0}$ is Abel summable in $\mathbb{C}$ to the limit s, then $\{s_n\}$ is convergent to the limit s.

PROOF. Let $\{r_j\}_{j\geq 0}$ be a sequence increasing to the limit 1. Since $a_n r_j^n \geq 0$ is nonnegative and since it is monotone increasing in j for each n, Corollary 1.14 applies and gives $\lim_j \sum_{n=0}^\infty a_n r_j^n = \sum_{n=0}^\infty \lim_j a_n r_j^n$, the limits existing in $\mathbb{R}^*$. The left side is the (finite) limit s of the Abel sums, and the right side is $\lim s_n$, which Corollary 1.14 is asserting exists.

EXAMPLE. The binomial series expansion in Section 7 shows, for any complex p, that $(1-r)^p$ is given for $-1 < r < 1$ by the absolutely convergent series

$$(1-r)^p = 1 + \sum_{n=1}^\infty (-1)^n \frac{p(p-1)\cdots(p-n+1)}{n!}\, r^n.$$

For p real with $0 < p < 1$, inspection shows that all the coefficients in the sum on the right are ≤ 0. Therefore

$$1-(1-r)^p = \sum_{n=1}^\infty (-1)^{n+1} \frac{p(p-1)\cdots(p-n+1)}{n!}\, r^n \qquad (*)$$

has all coefficients ≥ 0 if $0 < p < 1$. For $0 \leq r < 1$, the sum of the series is $1-(1-r)^p$ and is ≥ 0. The fact that $\lim_{r\uparrow 1}[1-(1-r)^p] = 1$ means that the sequence of partial sums $s_k = \sum_{k=1}^\infty (-1)^{k+1} \frac{p(p-1)\cdots(p-k+1)}{k!}$ is Abel summable to 1. Proposition 1.49 shows that the series $(*)$ is convergent at $r = 1$, and the Weierstrass M test shows that $(*)$ converges uniformly for $-1 \leq r \leq 1$ to $1-(1-r)^p$. If we now take $p = \frac{1}{2}$, we have

$$\begin{aligned}
(1-r)^{1/2} &= 1 - \sum_{n=1}^\infty (-1)^{n+1} \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2})\cdots(\frac{3}{2}-n)}{n!}\, r^n \\
&= \sum_{n=1}^\infty (-1)^{n+1} \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2})\cdots(\frac{3}{2}-n)}{n!} \\
&\qquad - \sum_{n=1}^\infty (-1)^{n+1} \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2})\cdots(\frac{3}{2}-n)}{n!}\, r^n \\
&= \sum_{n=1}^\infty (-1)^{n+1} \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2})\cdots(\frac{3}{2}-n)}{n!}\, (1-r^n),
\end{aligned}$$

the series on the right being uniformly convergent for $-1 \leq r \leq 1$. Putting $r = 1-x^2$ therefore gives

$$|x| = \sqrt{x^2} = \sum_{n=1}^\infty (-1)^{n+1} \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2})\cdots(\frac{3}{2}-n)}{n!} \big(1-(1-x^2)^n\big),$$

the series on the right being uniformly convergent for $-1 \le x \le 1$. Consequently $|x|$ is the uniform limit of a sequence of polynomials on $[-1, 1]$, and all these polynomials are in fact 0 at $x = 0$.

Proposition 1.50. Let $\{a_n\}_{n\ge 0}$ be a complex sequence, and let $\{s_n\}_{n\ge 0}$ be the sequence of partial sums with $s_n = \sum_{k=0}^n a_k$. If $\{s_n\}$ is Cesàro summable to the limit s in $\mathbb{C}$ and if the sequence $\{na_n\}$ is bounded, then $\{s_n\}$ is convergent and the limit is s. The rate of convergence depends only on the bound for $\{na_n\}$ and the rate of convergence of the Cesàro sums.

REMARK. In our application in Chapter VI to pointwise convergence of Fourier series, the sequence of partial sums will be of the form $\{s_n(x)\}$, depending on a parameter x, and the statement about the rate of convergence will enable us to see that the convergence of $\{s_n(x)\}$ is uniform in x under suitable hypotheses.

PROOF. Let $\{s_n\}$ be the sequence of partial sums of $\{a_n\}$, and choose M such that $|na_n| \le M$ for all n. The first step is to establish a useful formula for $s_n - \sigma_n$. Let m be any integer with $0 \le m < n$. We start from the trivial identity $-(n-m)\sigma_n = (m+1)\sigma_n - (n+1)\sigma_n$, add $(n-m)s_n$ to both sides, and regroup as

$$\begin{aligned}(n-m)(s_n - \sigma_n) &= (m+1)\sigma_n - s_0 - \cdots - s_m + (n-m)s_n - s_{m+1} - \cdots - s_n \\ &= (m+1)(\sigma_n - \sigma_m) + \sum_{j=m+1}^{n} (s_n - s_j).\end{aligned}$$

Dividing by $(n-m)$ yields

$$s_n - \sigma_n = \frac{m+1}{n-m}(\sigma_n - \sigma_m) + \frac{1}{n-m}\sum_{j=m+1}^{n}(s_n - s_j),$$

which is the identity from which the estimates begin.

For $m + 1 \le j \le n$, we have

$$\begin{aligned}|s_n - s_j| &\le |a_n| + |a_{n-1}| + \cdots + |a_{j+1}| \le \frac{M}{n} + \frac{M}{n-1} + \cdots + \frac{M}{j+1} \\ &\le \frac{M}{j+1} + \frac{M}{j+1} + \cdots + \frac{M}{j+1} = \frac{(n-j)M}{j+1} \le \frac{(n-m-1)M}{m+2}.\end{aligned}$$

Substituting into our identity yields

$$|s_n - \sigma_n| \le \frac{m+1}{n-m}\big|\sigma_n - \sigma_m\big| + \frac{(n-m-1)M}{m+2}.$$

Let $\epsilon > 0$ be given, and choose N such that $|\sigma_k - s| \le \epsilon^2$ whenever $k \ge N$. We may assume that $\epsilon < \frac{1}{2}$ and $N \ge 4$. With ϵ fixed and with n fixed to be $\ge 2N$, define m to be the unique integer with

$$m \le \frac{n-\epsilon}{1+\epsilon} < m+1.$$

Then $0 \le m < n$, and our inequality for $|s_n - \sigma_n|$ applies. From the left inequality $m \le \frac{n-\epsilon}{1+\epsilon}$ defining m, we obtain $m + m\epsilon \le n - \epsilon$ and hence $(m+1)\epsilon \le n - m$ and $\frac{m+1}{n-m} \le \epsilon^{-1}$. From the right inequality $\frac{n-\epsilon}{1+\epsilon} < m+1$ defining m, we obtain $n - \epsilon < m + 1 + \epsilon m + \epsilon$ and hence $n - m - 1 < \epsilon(m+2)$ and $\frac{n-m-1}{m+2} < \epsilon$. Thus our main inequality becomes

$$|s_n - \sigma_n| \le \epsilon^{-1}|\sigma_n - \sigma_m| + M\epsilon.$$

To handle σ_m, we need to bound m below. We have seen that $n - m - 1 < \epsilon(m+2)$, and we have assumed that $\epsilon < \frac{1}{2}$. Then $n - m - 1 < \frac{1}{2}(m+2)$, and this simplifies to $m > \frac{2n}{3} - \frac{4}{3}$, which is $\ge \frac{n}{2}$ if $n \ge 8$, thus certainly if $N \ge 4$. In other words, $N \ge 4$ and $n \ge 2N$ makes $m \ge \frac{n}{2} \ge N$. Therefore $|\sigma_m - s| < \epsilon^2$, and $|\sigma_n - \sigma_m| < 2\epsilon^2$. Substituting into our main inequality, we obtain

$$|s_n - \sigma_n| < \epsilon^{-1} 2\epsilon^2 + M\epsilon = (M+2)\epsilon.$$

Since ϵ is arbitrary, the proof is complete.

9. Weierstrass Approximation Theorem

We saw as an application of Proposition 1.49 that the function $|x|$ on $[-1, 1]$ is the uniform limit of an explicit sequence $\{P_n\}$ of polynomials with $P_n(0) = 0$. This is a special case of a theorem of Weierstrass that any continuous complex-valued function on a bounded interval is the uniform limit of polynomials on the interval.

The device for proving the Weierstrass theorem for a general continuous complex-valued function is to construct the approximating polynomials as the result of a smoothing process, known as the use of an "approximate identity." The idea of an approximate identity is an important one in analysis and will occur several times in this book. If f is the given function, the smoothing is achieved by "convolution"

$$\int f(x-t)\varphi(t)\,dt$$

of f with some function φ, the integrals being taken over some particular intervals. The resulting function of x from the convolution turns out to be as "smooth" as the smoother of f and φ. In the case of the Weierstrass theorem, the function

φ will be a polynomial, and we shall arrange parameters so that the convolution will automatically be a polynomial.

To see how a polynomial $\int f(x-t)\varphi(t)\,dt$ might approximate f, one can think of φ as some kind of mass distribution; the mass is all nonnegative if $\varphi \geq 0$. The integration produces a function of x that is the "average" of translates $x \mapsto f(x-t)$ of f, the average being computed according to the mass distribution φ. If φ has total mass 1, i.e., total integral 1, and most of the mass is concentrated near $t=0$, then f is being replaced essentially by an average of its translates, most of the translates being rather close to f, and we can expect the result to be close to f.

For the Weierstrass theorem, we use a single starting φ_1 at stage 1, namely $c_1(1-x^2)$ on $[-1,1]$ with c_1 chosen so that the total integral is 1. The graph of φ_1 is a familiar inverted parabola, with the appearance of a bump centered at the origin. The function at stage n is $c_n(1-x^2)^n$, with c_n chosen so that the total integral is 1. Graphs for $n=3$ and $n=30$ appear in Figure 1.1. The bump near the origin appears to be more pronounced at n increases, and what we need to do is to translate the above motivation into a proof.

Lemma 1.51. If c_n is chosen so that $c_n \int_{-1}^{1} (1-x^2)^n\,dx = 1$, then $c_n \leq e\sqrt{n}$ for n sufficiently large.

PROOF. We have

$$c_n^{-1} = \int_{-1}^{1} (1-x^2)^n\,dx \geq \int_{-1/\sqrt{n}}^{1/\sqrt{n}} (1-x^2)\,dx = 2\int_0^{1/\sqrt{n}} (1-x^2)^n\,dx$$
$$\geq 2\int_0^{1/\sqrt{n}} (1-\tfrac{1}{n})^n\,dx = 2(1-\tfrac{1}{n})^n\big/\sqrt{n}.$$

Since $(1-\frac{1}{n})^n \to e^{-1}$, we have $(1-\frac{1}{n})^n \geq \frac{1}{2}e^{-1}$ for n large enough (actually for $n \geq 2$). Therefore $c_n^{-1} \geq e^{-1}/\sqrt{n}$ for n large enough, and hence $c_n \leq e\sqrt{n}$ for n large enough. This proves the lemma.

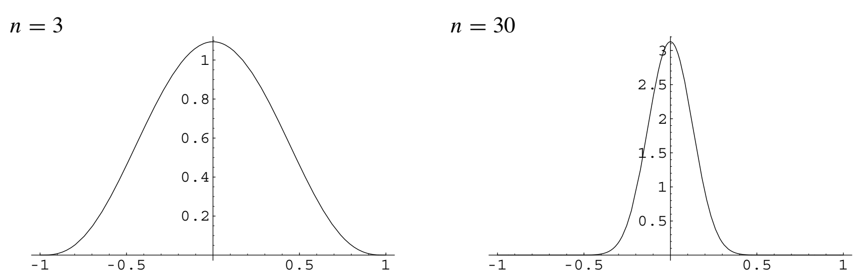


FIGURE 1.1. Approximate identity. Graphs of $c_n \int_{-1}^{1} (1-x^2)^n\,dx$ for for $n=3$ and $n=30$ with different scales used on the vertical axes.

Let $\varphi_n(x) = c_n(1-x^2)^n$ on $[-1, 1]$, with c_n as in the lemma. The polynomials φ_n have the following properties:

(i) $\varphi_n(x) \geq 0$,
(ii) $\int_{-1}^{1} \varphi_n(x)\,dx = 1$,
(iii) for any $\delta > 0$, $\sup\{\varphi_n(x) \mid \delta \leq x \leq 1\}$ tends to 0 as n tends to infinity.

Lemma 1.51 is used to verify (iii): the quantity

$$\sup\{\varphi_n(x) \mid \delta \leq x \leq 1\} = c_n(1-\delta^2)^n$$

tends to 0 because $\lim_n \sqrt{n}(1-\delta^2)^n = 0$. A function with the above three properties will be called an **approximate identity** on $[-1, 1]$.

Theorem 1.52 (Weierstrass Approximation Theorem). Any complex-valued continuous function on a bounded interval $[a, b]$ is the uniform limit of a sequence of polynomials.

PROOF. In order to arrange for the convolution to be a polynomial, we need to make some preliminary normalizations. Approximating $f(x)$ on $[a, b]$ by $P(x)$ uniformly within ϵ is the same as approximating $f(x+a)$ on $[0, b-a]$ by $P(x+a)$ uniformly within ϵ, and approximating $g(x)$ on $[0, c]$ uniformly by $Q(x)$ is the same as approximating $g(cx)$ uniformly by $Q(cx)$. Thus we may assume without loss of generality that $[a, b] = [0, 1]$.

If $h : [0, 1] \to \mathbb{C}$ is continuous and if r is the function defined by $r(x) = h(x) - h(0) - [h(1) - h(0)]x$, then r is continuous with $r(0) = r(1) = 0$. Approximating $h(x)$ on $[0, 1]$ uniformly by $R(x)$ is the same as approximating $r(x)$ on $[0, 1]$ uniformly by $R(x) - h(0) - [h(1) - h(0)]x$. Thus we may assume without loss of generality that the function to be approximated has value 0 at 0 and 1.

Let $f : [0, 1] \to \mathbb{C}$ be a given continuous function with $f(0) = f(1) = 0$; the function f is uniformly continuous by Theorem 1.10. We extend f to the whole line by making it be 0 outside $[0, 1]$, and the uniform continuity is maintained. Now let φ_n be the polynomial above, and put $P_n(x) = \int_{-1}^{1} f(x-t)\varphi_n(t)\,dt$.

Let us see that P_n is a polynomial. By our definition of the extended f, the integrand is 0 for a particular $x \in [0, 1]$ unless t is in $[x-1, x]$ as well as $[-1, 1]$. We change variables, replacing t by $s + x$ and making use of Theorem 1.34, and the integral becomes $P_n(x) = \int f(-s)\varphi_n(s+x)\,ds$, the integral being taken for s in $[-1, 0] \cap [-1-x, 1-x]$. Since x is in $[0, 1]$, the condition on s is that s is in $[-1, 0]$. Thus $P_n(x) = \int_{-1}^{0} f(-s)\varphi_n(s+x)\,ds$. In this integral, $\varphi_n(x)$ is a linear combination of monomials x^k, and x^k itself contributes $\int_{-1}^{0} f(-s)(x+s)^k\,ds$, which expands out to be a polynomial in x. Thus $P_n(x)$ is a polynomial in x.

By property (ii) of φ_n, we have

$$P_n(x) - f(x) = \int_{-1}^{1} f(x-t)\varphi_n(t)\,dt - f(x) = \int_{-1}^{1} [f(x-t) - f(x)]\,\varphi_n(t)\,dt.$$

Then property (i) gives

$$\begin{aligned} |P_n(x) - f(x)| &\le \int_{-1}^{1} |f(x-t) - f(x)|\,\varphi_n(t)\,dt \\ &= \int_{-\delta}^{\delta} |f(x-t) - f(x)|\,\varphi_n(t)\,dt + \Big(\int_{-1}^{-\delta} + \int_{\delta}^{1}\Big)|f(x-t) - f(x)|\,\varphi_n(t)\,dt, \end{aligned}$$

and two further uses of property (ii) show that this is

$$\le \sup_{|t|\le\delta} |f(x-t) - f(x)| + 4\Big(\sup_{y\in[0,1]} |f(y)|\Big)\Big(\sup_{\delta\le|t|\le 1} \varphi_n(t)\Big).$$

Given $\epsilon > 0$, we choose some δ of uniform continuity for f and ϵ, and then the first term is $\le \epsilon$. With δ fixed, we use property (iii) of φ_n and the boundedness of f, given by Theorem 1.11, to produce an integer N such that the second term is $< \epsilon$ for $n \ge N$. Then $n \ge N$ implies that the displayed expression is $< 2\epsilon$. Since ϵ is arbitrary, P_n converges uniformly to f.

10. Fourier Series

A **trigonometric series** is a series of the form $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ with complex coefficients. The individual terms of the series thus form a doubly infinite sequence, but the sequence of partial sums is always understood to be the sequence $\{s_N\}_{N=0}^{\infty}$ with $s_N(x) = \sum_{n=-N}^{N} c_n e^{inx}$. Such a series may also be written as

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

by putting

$$\left.\begin{aligned} e^{inx} &= \cos nx + i \sin nx \\ e^{-inx} &= \cos nx - i \sin nx \end{aligned}\right\} \quad \text{for } n > 0,$$

$$c_0 = \tfrac{1}{2}a_0, \quad c_n = \tfrac{1}{2}(a_n - ib_n), \quad \text{and} \quad c_{-n} = \tfrac{1}{2}(a_n + ib_n) \quad \text{for } n > 0.$$

Historically the notation with the a_n's and b_n's was introduced first, but the use of complex exponentials has become quite common. Nowadays the notation with

a_n's and b_n's tends to be used only when a function f under investigation is real-valued or when all the cosine terms are absent (i.e., f is even) or all the sine terms are absent (i.e., f is odd).

Power series enable us to enlarge our repertory of explicit functions, and the same thing is true of trigonometric series. Just as the coefficients of a power series whose sum is a function f have to be those arising from Taylor's formula for f, the coefficients of a trigonometric series formed from a function have to arise from specific formulas. Let us run through the relevant formal computation: First we observe that the partial sums have to be periodic with period 2π. The question then is the extent to which a complex-valued periodic function f on the real line can be given by a trigonometric series. Suppose that

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

Multiply by e^{-ikx} and integrate to get

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{n=-\infty}^{\infty} c_n e^{inx}e^{-ikx}\,dx.$$

If we can interchange the order of the integration and the infinite sum, e.g., if the trigonometric series is uniformly convergent to $f(x)$, the right side is

$$= \sum_{n=-\infty}^{\infty} c_n \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-ikx}\,dx = \sum_{n=-\infty}^{\infty} c_n \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-k)x}\,dx = c_k$$

because

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{imx}\,dx = \begin{cases} 1 & \text{if } m = 0, \\ 0 & \text{if } m \neq 0. \end{cases}$$

Let f be Riemann integrable on $[-\pi, \pi]$, and regard f as periodic on $\mathbb{R}$. The trigonometric series $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ with

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$$

is called the **Fourier series** of f. We write

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx} \qquad \text{and} \qquad s_N(f;x) = \sum_{n=-N}^{N} c_n e^{inx}.$$

The numbers c_n are the **Fourier coefficients** of f, and the functions $s_N(f;x)$ are the partial sums of the Fourier series. The symbol $\sim$ is to be read as "has Fourier series," nothing more, at least initially. The formulas for the coefficients when the Fourier series is written with sines and cosines are

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx \qquad \text{for } n \geq 0,$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx \qquad \text{for } n \geq 1.$$

In applications one encounters periodic functions of periods other than 2π. If f is periodic of period $2l$, then the Fourier series of f is $f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/l}$ with $c_n = (2l)^{-1}\int_{-l}^{l} f(x)e^{-in\pi x/l}\,dx$. The formula for the series written with sines and cosines is $f(x) \sim a_0/2 + \sum_{n=1}^{\infty}(a_n\cos(n\pi x/l) + b_n\sin(n\pi x/l))$ with $a_n = l^{-1}\int_{-l}^{l} f(x)\cos(n\pi x/l)\,dx$ and $b_n = l^{-1}\int_{-l}^{l} f(x)\sin(n\pi x/l)\,dx$. In the present section of the text, we shall assume that our periodic functions have period 2π.

The result implicit in the formal computation above is that if $f(x)$ is the sum of a uniformly convergent trigonometric series, then the trigonometric series is the Fourier series of f, by Theorem 1.31.

We ask two questions: When does a general Fourier series converge? If the Fourier series converges, to what extent does the sum represent f? We begin with an illuminating example that brings together a number of techniques from this chapter.

EXAMPLE. As in the example following Theorem 1.48, we have

$$\log\Big(\frac{1}{1-x}\Big) = x + \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 + \cdots \qquad \text{for } -1 < x < 1.$$

We would like to extend this identity to complex z with $|z| < 1$ but do not want to attack the problem of making sense out of log as a function of a complex variable. What we do is apply exp to both sides and obtain an identity for which both sides make sense when the real x is replaced by a complex z:

$$\exp\big(z + \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 + \cdots\big) = \frac{1}{1-z} \qquad \text{for } |z| < 1.$$

In fact, this identity is valid for z complex with $|z| < 1$, and Problems 30–35 at the end of the chapter lead to a proof of it. Corollary 1.45 allows us to write $z = re^{i\theta}$ and $z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \cdots = \rho e^{i\varphi}$. Equating real and imaginary parts of the latter equation gives us

$$\rho\cos\varphi = \sum_{n=1}^{\infty}\frac{r^n\cos n\theta}{n} \qquad \text{and} \qquad \rho\sin\varphi = \sum_{n=1}^{\infty}\frac{r^n\sin n\theta}{n}.$$

We shall compute the left sides of these displayed equations in another way. We have

$$e^{\rho\cos\varphi}e^{i\rho\sin\varphi} = \exp(\rho\cos\varphi + i\rho\sin\varphi) = \exp(\rho e^{i\varphi}) = (1-z)^{-1}$$

and therefore also $e^{\rho\cos\varphi}e^{-i\rho\sin\varphi} = (1-\bar z)^{-1}$. Thus

$$e^{2\rho\cos\varphi} = (1-z)^{-1}(1-\bar z)^{-1} = (1-re^{i\theta})^{-1}(1-re^{-i\theta})^{-1} = (1-2r\cos\theta+r^2)^{-1}.$$

Taking log of both sides gives $2\rho\cos\varphi = \log\big((1-2r\cos\theta+r^2)^{-1}\big)$, and thus we have

$$\tfrac{1}{2}\log\Big(\frac{1}{1-2r\cos\theta+r^2}\Big) = \sum_{n=1}^{\infty}\frac{r^n\cos n\theta}{n}. \tag{$*$}$$

Handling $\rho\sin\varphi$ is a little harder. From $e^{\rho\cos\varphi}e^{i\rho\sin\varphi} = (1-z)^{-1}$, we have $e^{i\rho\sin\varphi} = (1-z)^{-1}/|1-z|^{-1} = (1-\bar z)/|1-z| = \frac{1-r\cos\theta}{|1-z|} + i\frac{r\sin\theta}{|1-z|}$, and hence

$$\cos(\rho\sin\varphi) = (1-r\cos\theta)/|1-z| \quad\text{and}\quad \sin(\rho\sin\varphi) = (r\sin\theta)/|1-z|.$$

Thus $\tan(\rho\sin\varphi) = r\sin\theta/(1-r\cos\theta)$. Since $1-r\cos\theta$ is >0, $\cos(\rho\sin\varphi)$ is >0, and $\rho\sin\varphi = \arctan\big((r\sin\theta)/(1-r\cos\theta)\big) + 2\pi N(r,\theta)$ for some integer $N(r,\theta)$ depending on r and θ. Hence

$$\arctan\big((r\sin\theta)/(1-r\cos\theta)\big) + 2\pi N(r,\theta) = \sum_{n=1}^{\infty}\frac{r^n\sin n\theta}{n}.$$

For fixed r, the first term on the left is continuous in θ, and the series on the right is uniformly convergent by the Weierstrass M test. By Theorem 1.21 the right side is continuous in θ. Thus $N(r,\theta)$ is continuous in θ for fixed r; since $N(r,0)=0$, $N(r,\theta)=0$ for all r and θ. We conclude that

$$\arctan\Big(\frac{r\sin\theta}{1-r\cos\theta}\Big) = \sum_{n=1}^{\infty}\frac{r^n\sin n\theta}{n}. \tag{$**$}$$

Problem 15 at the end of the chapter observes that the partial sums $\sum_{n=1}^{N}\cos n\theta$ and $\sum_{n=1}^{N}\sin n\theta$ are uniformly bounded on any set $\epsilon \le \theta < \pi-\epsilon$ if $\epsilon > 0$. Corollary 1.19 therefore shows that the series

$$\sum_{n=1}^{\infty}\frac{\cos n\theta}{n} \qquad\text{and}\qquad \sum_{n=1}^{\infty}\frac{\sin n\theta}{n}$$

are uniformly convergent for $\epsilon \le \theta < \pi-\epsilon$ if $\epsilon > 0$. Abel's Theorem (Theorem 1.48) shows that each of these series is therefore Abel summable with the same limit. We can tell what the latter limits are from $(*)$ and $(**)$, and thus we conclude that

$$\tfrac{1}{2}\log\left(\frac{1}{2-2\cos\theta}\right)=\sum_{n=1}^{\infty}\frac{\cos n\theta}{n}$$

and

$$\arctan\left(\frac{\sin\theta}{1-\cos\theta}\right)=\sum_{n=1}^{\infty}\frac{\sin n\theta}{n},$$

The sum of the series with the cosine terms is unbounded near $\theta=0$, and Riemann integration is not meaningful with it. We shall not be able to analyze this series further until we can treat the left side in Chapter VI by means of Lebesgue integration. The sum of the series with the sine terms is written in a way that stresses its periodicity. On the interval $[-\pi,\pi]$, we can rewrite its left side as $\frac{1}{2}(-\pi-\theta)$ for $-\pi\le\theta<0$, 0 for $\theta=0$, and $\frac{1}{2}(\pi-\theta)$ for $0<\theta\le\pi$. The expression for the left side is nicer on the interval $(0,2\pi)$, and there we have

$$\tfrac{1}{2}(\pi-\theta)=\sum_{n=1}^{\infty}\frac{\sin n\theta}{n}\qquad\text{for } 0<\theta<2\pi.$$

The function $\frac{1}{2}(\pi-\theta)$ is bounded on $(0,2\pi)$, and we can readily compute its Fourier coefficients from the formula $b_n=\pi^{-1}\int_0^{2\pi}\frac{1}{2}(\pi-\theta)\sin n\theta\,d\theta$, using integration by parts (Corollary 1.33). The result is that $b_n=1/n$. Hence the displayed series is the Fourier series. Graphs of some of the partial sums appear in Figure 1.2.

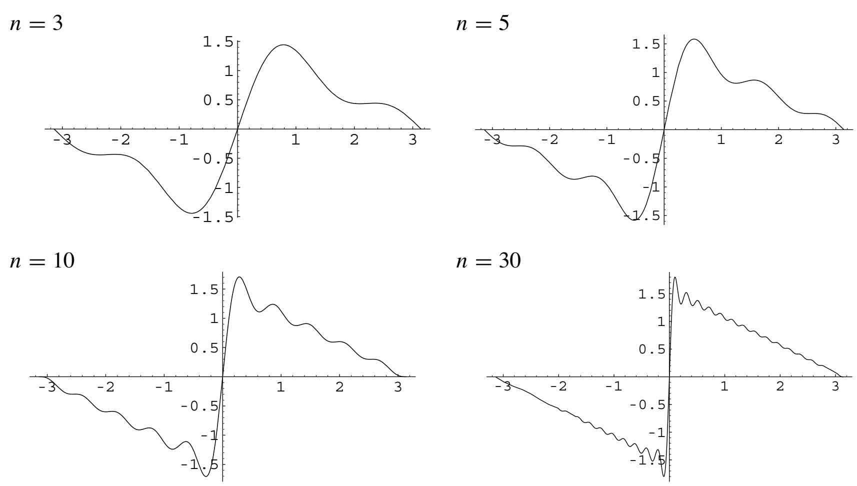


FIGURE 1.2. Fourier series of sawtooth function. Graphs of $\sum_{n=1}^{N}(\sin nx)/n$ for $n=3,5,10,30$.

The sawtooth function in the above example has a discontinuity, and yet its Fourier series converges to it pointwise. The recognition of the remarkable potential that Fourier series have for representing discontinuous functions dates to Joseph Fourier himself and caused many of Fourier's contemporaries to doubt the validity of his work.

Although the above Fourier series converges to the function, it cannot do so uniformly, as a consequence of Theorem 1.21. In any such situation the Fourier coefficients cannot decrease rapidly, and a decrease of order $1/n$ is the best that one gets for a nice function with a jump discontinuity.

This example points to a general heuristic principle contrasting how power series and trigonometric series behave: whereas Taylor series converge very rapidly and may not converge to the function, Fourier series are inclined to converge rather slowly and they are more likely to converge to the function.

We come to convergence results in a moment. First we establish some elementary properties of them. Taking the absolute value of c_n in the definition of Fourier coefficient, we obtain the trivial bound $|c_n| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|\,dx$.

Theorem 1.53. Let f be in $\mathcal{R}[-\pi,\pi]$. Among all choices of $d_{-N},\dots,d_N$, the expression

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Big|f(x)-\sum_{n=-N}^{N} d_n e^{inx}\Big|^2\,dx$$

is minimized uniquely by choosing d_n, for all n with $|n|\le N$, to be the Fourier coefficient $c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$. The minimum value is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)|^2\,dx - \sum_{n=-N}^{N}|c_n|^2.$$

PROOF. Put $d_n = c_n + \varepsilon_n$. Then

$$\begin{aligned}
&\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}\big|f(x)-\sum_{n=-N}^{N} d_n e^{inx}\big|^2\,dx\\
&\qquad= \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}|f(x)|^2\,dx - \tfrac{1}{2\pi}2\operatorname{Re}\sum_{n=-N}^{N}\overline{d_n}\int_{-\pi}^{\pi}f(x)e^{-inx}\,dx\\
&\qquad\quad+\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}\sum_{m,n=-N}^{N} d_m\overline{d_n}e^{i(m-n)x}\,dx\\
&\qquad= \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}|f(x)|^2\,dx - 2\operatorname{Re}\sum_{n=-N}^{N}c_n\overline{d_n}+\sum_{n=-N}^{N}|d_n|^2\\
&\qquad= \big(\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}|f(x)|^2\,dx\big)-\big(2\sum_{n=-N}^{N}|c_n|^2+2\operatorname{Re}\sum_{n=-N}^{N}c_n\overline{\varepsilon_n}\big)\\
&\qquad\quad+\big(\textstyle\sum_{n=-N}^{N}|c_n|^2+2\operatorname{Re}\sum_{n=-N}^{N}c_n\overline{\varepsilon_n}+\sum_{n=-N}^{N}|\varepsilon_n|^2\big)\\
&\qquad= \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}|f(x)|^2-\sum_{n=-N}^{N}|c_n|^2+\sum_{n=-N}^{N}|\varepsilon_n|^2.
\end{aligned}$$

The result follows.

Corollary 1.54 (Bessel's inequality). Let f be in $\mathcal{R}[-\pi, \pi]$, and let $f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$. Then

$$\sum_{n=-\infty}^{\infty} |c_n|^2 \le \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx.$$

In particular, $\sum_{n=-\infty}^{\infty} |c_n|^2$ is finite.

REMARK. In terms of the coefficients a_n and b_n, the corresponding result is

$$\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} \big(|a_n|^2 + |b_n|^2\big) \le \frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx.$$

PROOF. The theorem shows that the minimum value of a certain nonnegative quantity depending on N is $\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx - \sum_{n=-N}^{N} |c_n|^2$. Thus, for any N, $\sum_{n=-N}^{N} |c_n|^2 \le \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx$. Letting N tend to infinity, we obtain the corollary.

Corollary 1.55 (Riemann–Lebesgue Lemma). If f is in $\mathcal{R}[-\pi, \pi]$ and has Fourier coefficients $\{c_n\}_{n=-\infty}^{\infty}$, then $\lim_{|n| \to \infty} c_n = 0$.

REMARK. This improves on the inequality $|c_n| \le \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)| \, dx$ observed above, which shows, by means of an explicit estimate, that $\{c_n\}$ is a bounded sequence.

PROOF. This is immediate from Corollary 1.54.

We now turn to convergence results. First it is necessary to clarify terms like "continuous" and "differentiable" in the context of Fourier series of functions on $[-\pi, \pi]$. Each term of a Fourier series is defined on all of $\mathbb{R}$ and is periodic with period 2π and is really given as the restriction to $[-\pi, \pi]$ of this periodic function. Thus it makes sense to regard a general function in the same way if one wants to form its Fourier series: a function f is extended to all of $\mathbb{R}$ so as to be periodic with period 2π, and if we consider f on $[-\pi, \pi]$, it is really the restriction to $[-\pi, \pi]$ that we are considering.

In particular, it makes sense to insist that $f(-\pi) = f(\pi)$; if f does not have this property initially, one or both of these endpoint values will have to be adjusted, but that adjustment will not affect any Fourier coefficients. Similarly continuity of f will refer to continuity of the extended function on all of $\mathbb{R}$, and similarly for differentiability.

That being said, let us take up the matter of integration by parts for the functions we are considering. The scope of integration by parts in Corollary 1.33 was limited to a pair of functions f and g that have a continuous first derivative. In the context of Fourier series, it is the periodic extensions that are to have these properties, and then the integration-by-parts formula simplifies. Namely,

$$\begin{aligned}\int_{-\pi}^{\pi} f(x)g'(x)\,dx &= \Big[f(x)g(x)\Big]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f'(x)g(x)\,dx \\ &= -\int_{-\pi}^{\pi} f'(x)g(x)\,dx,\end{aligned}$$

i.e., the integrated term drops out because of the assumed periodicity.

The simplest convergence result for Fourier series is that a periodic function (of period 2π) with two continuous derivatives has a uniformly convergent Fourier series. To prove this, we take $n \neq 0$ and use the above integration-by-parts formula twice to obtain

$$\begin{aligned}c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx &= -\frac{1}{2\pi}\Big(\frac{1}{-in}\Big)\int_{-\pi}^{\pi} f'(x)e^{-inx}\,dx \\ &= \frac{1}{2\pi}\Big(\frac{1}{-in}\Big)^2\int_{-\pi}^{\pi} f''(x)e^{-inx}\,dx.\end{aligned}$$

Then $|c_n e^{inx}| = |c_n| \leq C/n^2$, where $C = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f''(x)|\,dx$, and the Fourier series converges uniformly by the Weierstrass M test. The argument does not say that the convergence is to f, but that fact will be proved in Theorem 1.57 below.

Adjusting the proof just given, we can prove a sharper convergence result.

Proposition 1.56 . If f is periodic (of period 2π) and has one continuous derivative, then the Fourier series of f converges uniformly.

PROOF. As in the above argument, $c_n = -\frac{1}{2\pi}\big(\frac{1}{-in}\big)\int_{-\pi}^{\pi} f'(x)e^{-inx}\,dx$, and this equals $\frac{1}{in}d_n$, where d_n is the n^{th} Fourier coefficient of the continuous function f'. In the computation that follows, we use the classical Schwarz inequality (as in Section A5 of the appendix) for finite sums and pass to the limit in order to get the first inequality, and then we use Bessel's inequality (Corollary 1.54) to get the second inequality:

$$\begin{aligned}\sum_{n\neq 0}|c_n| = \sum_{n\neq 0}|inc_n|\,\frac{1}{|n|} = \sum_{n\neq 0}\frac{1}{|n|}\,|d_n| &\leq \Big(\sum_{n\neq 0}\frac{1}{n^2}\Big)^{1/2}\Big(\sum_{n\neq 0}|d_n|^2\Big)^{1/2} \\ &\leq \Big(\sum_{n\neq 0}\frac{1}{n^2}\Big)^{1/2}\Big(\frac{1}{2\pi}\int_{-\pi}^{\pi}|f'(x)|^2\,dx\Big)^{1/2}.\end{aligned}$$

The right side is finite, and the proposition follows from the Weierstrass M test.

The fact that the convergence in Proposition 1.56 is actually to f will follow from Dini's test, which is Theorem 1.57 below. We first derive some simple formulas. The **Dirichlet kernel** is the periodic function of period 2π defined by

$$D_N(x) = \sum_{n=-N}^{N} e^{inx} = \frac{\sin\big((N+\frac{1}{2})x\big)}{\sin\frac{1}{2}x},$$

the second equality following from the formula for the sum of a geometric series. For a periodic function f of period 2π, the partial sums of the Fourier series of f are given by

$$\begin{aligned} s_N(f;x) &= \sum_{n=-N}^{N} \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt\Big)e^{inx} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\sum_{n=-N}^{N} e^{in(x-t)}\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_N(x-t)\,dt \\ &= \frac{1}{2\pi}\int_{x-\pi}^{x+\pi} f(x-s)D_N(s)\,ds \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)\,dt, \end{aligned}$$

the last two steps following from the changes of variables $t \mapsto x+s$ (Theorem 1.34) and $s \mapsto -s$ (Proposition 1.30h) and from the periodicity of f and D_N.

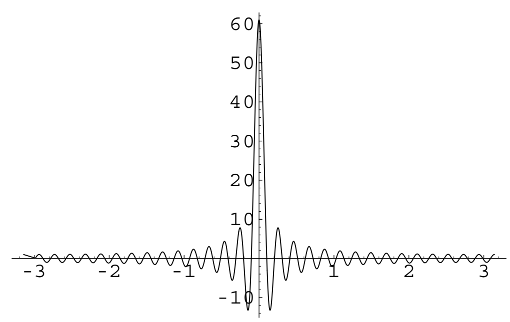


FIGURE 1.3. Dirichlet kernel. Graph of D_N for $N = 30$.

This is the kind of convolution integral that occurred in the previous section. Term-by-term integration shows that $\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(x)\,dx = 1$. However, D_N is not an approximate identity, not being everywhere ≥ 0. Figure 1.3 shows the graph of D_N for $N = 30$. Although $D_N(x)$ looks small in the graph away from $x = 0$, it is small only as a percentage of $D_N(0)$; $D_N(x)$ does not have $\lim_N D_N(x)$ equal to 0 for $x \neq 0$. Thus $D_N(x)$ fails in a second way to be an approximate identity. The failure of D_N to be an approximate identity is what makes the subject of convergence of Fourier series so subtle.

Theorem 1.57 (Dini's test). Let $f : \mathbb{R} \to \mathbb{C}$ be periodic of period 2π and Riemann integrable on $[-\pi, \pi]$. Fix x in $[-\pi, \pi]$. If there are constants $\delta > 0$ and $M < +\infty$ such that

$$|f(x+t) - f(x)| \le M|t| \qquad \text{for } |t| < \delta,$$

then $\lim_N s_n(f; x) = f(x)$.

REMARK. This condition is satisfied if f is differentiable at x. Thus the convergence of the Fourier series in Proposition 1.56 is to the original function f. By contrast, the Dini condition is not satisfied at $x = 0$ for the continuous periodic extension of the function $f(x) = |x|^{1/2}$ defined on $(-\pi, \pi]$.

PROOF. With x fixed, let

$$g(t) = \begin{cases} \dfrac{f(x-t) - f(x)}{\sin t/2} & \text{for } 0 < |t| \le \pi, \\ 0 & \text{for } t = 0. \end{cases}$$

Proposition 1.30d shows that $(\sin t/2)^{-1}$ is Riemann integrable on $\epsilon \le |t| \le \pi$ for any $\epsilon > 0$, and hence so is $g(t)$. Since $g(t)$ is bounded near $t = 0$, Lemma 1.28 shows that $g(t)$ is Riemann integrable on $[-\pi, \pi]$. Since $\int_{-\pi}^{\pi} D_N(x)\,dx = 1$, we have

$$\begin{aligned} s_N(f;x) - f(x) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\,\frac{\sin\big((N+\tfrac12)t\big)}{\sin\tfrac12 t}\,dt - \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,\frac{\sin\big((N+\tfrac12)t\big)}{\sin\tfrac12 t}\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\sin\big((N+\tfrac12)t\big)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \big[g(t)\cos\tfrac{t}{2}\big]\sin Nt\,dt + \frac{1}{2\pi}\int_{-\pi}^{\pi} \big[g(t)\sin\tfrac{t}{2}\big]\cos Nt\,dt, \end{aligned}$$

and both terms on the right side tend to 0 with N by the Riemann–Lebesgue Lemma (Corollary 1.55).

Dini's test (Theorem 1.57) has implications for "localization" of the convergence of Fourier series. Suppose that $f = g$ on an open interval I, and suppose that the Fourier series of f converges to f on I. Then Dini's test shows that the Fourier series of $f - g$ converges to 0 on I, and hence the Fourier series of g converges to g on I. For example, f could be a function with a continuous derivative everywhere, and g could have discontinuities outside the open interval I. For f, the proof of Proposition 1.56 shows that $\sum |c_n| < +\infty$. But for g, the Fourier series cannot converge so rapidly because the sum of a uniformly convergent series of continuous functions has to be continuous. Thus the two series locally have the same sum, but their qualitative behavior is quite different.

Next let us address the question of the extent to which the Fourier series of f uniquely determines f. Our first result in this direction will be that if f and g are Riemann integrable and have the same respective Fourier coefficients, then $f(x) = g(x)$ at every point of continuity of both f and g. It may look as if some sharpening of Dini's test might apply just under the assumption of continuity of the function, and then this uniqueness result would be trivial. However, as we shall see in Chapter XII, the Fourier series of a continuous function need not converge to the function at particular points, and there can be no such sharpening of Dini's test. Instead, we shall handle the uniqueness question in a more indirect fashion.

The technique is to use an approximate identity, as in the proof of the Weierstrass Approximation Theorem in Section 9. Although the partial sums of the Fourier series of a continuous function need not converge at every point, the Cesàro sums do converge. To get at this fact, we shall examine the **Fejér kernel**

$$K_N(x) = \frac{1}{N+1} \sum_{n=0}^{N} D_n(x).$$

The N^{th} Cesàro sum of $s_n(f;x)$ is given by $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(x-t) f(t)\,dt$ because

$$\begin{aligned}\frac{1}{N+1}\sum_{n=0}^{N} s_n(f;x) &= \frac{1}{N+1}\sum_{n=0}^{N} \frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(x-t) f(t)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(x-t) f(t)\,dt.\end{aligned}$$

The remarkable fact is that the Fejér kernel is an approximate identity even though the Dirichlet kernel is not, and the result will be that the Cesàro sums of a Fourier series converge in every way that they have any hope of converging.

Lemma 1.58. The Fejér kernel is given by

$$K_N(x) = \frac{1}{N+1}\,\frac{1-\cos(N+1)x}{1-\cos x}.$$

PROOF. We show by induction on N that the values of $K_N(x)$ in the definition and in the lemma are equal. For $N = 0$, we have $K_0(x) = D_0(x) = 1 = \frac{1-\cos 1x}{1-\cos x}$ as required. Assume the equality for $N-1$. Then

$$\begin{aligned}
(N+1)K_N(x) &= \sum_{n=0}^{N} D_n(x) = NK_{N-1}(x) + D_N(x) \\
&= \frac{1-\cos Nx}{1-\cos x} + \frac{\sin\big((N+\frac{1}{2})x\big)}{\sin\frac{1}{2}x} \cdot \frac{\sin\frac{1}{2}x}{\sin\frac{1}{2}x} \qquad \text{by induction} \\
&= \frac{1-\cos Nx + 2\sin\big((N+\frac{1}{2})x\big)\sin\frac{1}{2}x}{1-\cos x} \\
&= \frac{1-\cos Nx - \big[\cos\big((N+\frac{1}{2})x+\frac{1}{2}x\big) - \cos\big((N+\frac{1}{2})x-\frac{1}{2}x\big)\big]}{1-\cos x} \\
&= \frac{1-\cos(N+1)x}{1-\cos x},
\end{aligned}$$

as required.

In line with the definition of approximate identity in Section 9, we are to show that $K_N(x)$ has the following properties:

(i) $K_N(x) \geq 0$,
(ii) $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(x)\,dx = 1$,
(iii) for any $\delta > 0$, $\sup_{\delta \leq |x| \leq \pi} K_N(x)$ tends to 0 as n tends to infinity.

Property (i) follows from the definition of $K_N(x)$, since $\cos x \leq 1$ everywhere; (ii) follows from the definition of $K_N(x)$ and the linearity of the integral, since $\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(x)\,dx = 1$ for all n; and (iii) follows from Lemma 1.58, since $1-\cos(N+1)x \leq 2$ everywhere and $1-\cos x \geq 1-\cos\delta$ if $\delta \leq |x| \leq \pi$.

Theorem 1.59 (Fejér's Theorem). Let $f : \mathbb{R} \to \mathbb{C}$ be periodic of period 2π and Riemann integrable on $[-\pi, \pi]$. If f is continuous at a point x_0 in $[-\pi, \pi]$, then

$$\lim_{N\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)K_N(x_0 - x)\,dx = f(x_0).$$

If f is uniformly continuous on a subset E of $[-\pi, \pi]$, then the convergence is uniform for x_0 in E.

PROOF. Choose M such that $|f(x)| \le M$ for all x. By (ii) and then (i),

$$\begin{aligned}
\Big|\frac{1}{2\pi}\int_{-\pi}^{\pi} & f(x)K_N(x_0-x)\,dx - f(x_0)\Big| \\
&= \Big|\frac{1}{2\pi}\int_{-\pi}^{\pi}[f(x)-f(x_0)]K_N(x_0-x)\,dx\Big| \\
&\le \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)-f(x_0)|K_N(x_0-x)\,dx \\
&\le \frac{1}{2\pi}\int_{|x-x_0|\le\delta}|f(x)-f(x_0)|K_N(x_0-x)\,dx \\
&\quad + \frac{1}{2\pi}\int_{\delta\le|x-x_0|\le\pi} 2M\Big(\sup_{\delta\le|t|\le\pi}K_N(t)\Big)\,dx \\
&\le \frac{1}{2\pi}\int_{|x-x_0|\le\delta}|f(x)-f(x_0)|K_N(x_0-x)\,dx + 2M\sup_{\delta\le|t|\le\pi}K_N(t).
\end{aligned}$$

Given $\epsilon > 0$, choose some δ for ϵ and continuity of f at x_0 or for ϵ and uniform continuity of f on E. In the first term on the right side, we then have $|f(x)-f(x_0)| \le \epsilon$ on the set where $|x-x_0| \le \delta$. Thus a second use of (i) shows that the above expression is

$$\le \epsilon + 2M\sup_{\delta\le|t|\le\pi}K_N(t).$$

With δ fixed, property (iii) shows that the right side is $< 2\epsilon$ if N is sufficiently large, and the theorem follows.

Corollary 1.60 (uniqueness theorem). Let $f : \mathbb{R} \to \mathbb{C}$ and $g : \mathbb{R} \to \mathbb{C}$ be periodic of period 2π and Riemann integrable on $[-\pi, \pi]$. If f and g have the same respective Fourier coefficients, then $f(x) = g(x)$ at every point of continuity of both f and g.

REMARK. The fact that f and g have the same Fourier coefficients means that $s_n(f;x) = s_n(g;x)$ for all n, hence that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(x-t)f(t)\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(x-t)g(t)\,dt$$

for all n. Then the same formula applies with D_n replaced by its Cesàro sums K_N.

PROOF. Apply Theorem 1.59 to $f-g$ at a point x_0 of continuity of both f and g.

Our second result about uniqueness will improve on Corollary 1.60, saying that any Riemann integrable function with all Fourier coefficients 0 is basically the 0 function—at least in the sense that any definite integral in which it is a factor of the integrand is 0. We shall prove this improved result as a consequence of Parseval's Theorem, which says that equality holds in Bessel's inequality. The proof of Parseval's Theorem will be preceded by an example and some lemmas.

Theorem 1.61 (Parseval's Theorem). Let $f : \mathbb{R} \to \mathbb{C}$ be periodic of period 2π and Riemann integrable on $[-\pi, \pi]$. If $f(x) \sim \sum_{-\infty}^{\infty} c_n e^{inx}$, then

$$\lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - s_N(f; x)|^2\, dx = 0$$

and

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2\, dx = \sum_{n=-\infty}^{\infty} |c_n|^2.$$

REMARK. In terms of the coefficients a_n and b_n, the corresponding result is

$$\frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2\, dx = \frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} \big(|a_n|^2 + |b_n|^2\big).$$

EXAMPLE. We saw near the beginning of this section that the periodic function f given by $f(x) = \frac{1}{2}(\pi - x)$ on $(0, 2\pi)$ has $f(x) \sim \sum\limits_{n=1}^{\infty} \dfrac{\sin nx}{n}$. The formulation of Parseval's Theorem as in the remark, but with the interval $(0, 2\pi)$ replacing the interval $(-\pi, \pi)$, says that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{\pi} \int_0^{2\pi} \big|\frac{1}{2}(\pi - x)\big|^2\, dx$. The right side is $= \frac{1}{4\pi} \int_{-\pi}^{\pi} x^2\, dx = \frac{2\pi^3/3}{4\pi} = \frac{\pi^2}{6}$. Thus

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

This formula was discovered by Euler by other means before the work of Fourier.

For the purposes of the lemmas and the proof of Parseval's Theorem, let us introduce a "Hermitian inner product"[3] on $\mathcal{R}[-\pi, \pi]$ by the definition

$$(f, g)_2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}\, dx,$$

[3] The term "Hermitian inner product" will be defined precisely in Section II.1. The form $(f, g)_2$ comes close to being one, but it fails to meet all the conditions because $(f, f)_2 = 0$ is possible without $f = 0$.

as well as a "norm" defined by

$$\|f\|_2 = (f, f)_2^{1/2} = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx\Big)^{1/2}$$

and a "distance function" defined by

$$d_2(f, g) = \|f - g\|_2 = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - g(x)|^2\,dx\Big)^{1/2}.$$

The role of the function d_2 will become clearer in Chapter II, where "distance functions" of this kind will be studied extensively.

Lemma 1.62. If f is in $\mathcal{R}[-\pi, \pi]$ and $\int_{-\pi}^{\pi} |f(x)|^2\,dx = 0$, then $\int_{-\pi}^{\pi} |f(x)|\,dx = 0$ and also $\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx = 0$ for all $g \in \mathcal{R}[-\pi, \pi]$.

PROOF. Write $M = \sup_{x\in[-\pi,\pi]} |f(x)|$, and let $\epsilon > 0$ be given. Choose a partition $P = \{x_i\}_{i=0}^n$ with $U(P, |f|^2) < \epsilon^3$, i.e.,

$$\sum_{i=1}^{n} \Big(\sup_{x\in[x_{i-1},x_i]} |f(x)|^2\Big)\Delta x_i \le \epsilon^3.$$

Divide the indices from 1 to n into two subsets, A and B, with

$$A = \Big\{i \mid \sup_{x\in[x_{i-1},x_i]} |f(x)| \ge \epsilon\Big\} \qquad \text{and} \qquad B = \Big\{i \;\Big|\; \sup_{x\in[x_{i-1},x_i]} |f(x)| < \epsilon\Big\}.$$

The sum of the contributions from indices $i \in A$ to $U\big(P, |f|^2\big)$ is $\ge \epsilon^2 \sum_{i\in A} \Delta x_i$, and thus $\sum_{i\in A} \Delta x_i \le \epsilon$. Hence $\sum_{i\in A} \big(\sup_{x\in[x_{i-1},x_i]} |f(x)|\big)\Delta x_i \le M\epsilon$. Also, $\sum_{i\in B} \big(\sup_{x\in[x_{i-1},x_i]} |f(x)|\big)\Delta x_i \le 2\pi\epsilon$. Therefore $U(P, |f|) \le (2\pi + M)\epsilon$. Since ϵ is arbitrary, $\int_{-\pi}^{\pi} |f(x)|\,dx = 0$. This proves the first conclusion.

For the second conclusion it follows from the boundedness of $|g|$, say by M', that $\big|\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx\big| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)||g(x)|\,dx \le \frac{M'}{2\pi}\int_{-\pi}^{\pi} |f(x)|\,dx = 0$.

Lemma 1.63 (Schwarz inequality). If f and g are in $\mathcal{R}[-\pi, \pi]$, then

$$|(f, g)_2| \le \|f\|_2\|g\|_2.$$

REMARK. Compare this result with the version of the Schwarz inequality in Section A5 of the appendix. This kind of inequality is put into a broader setting in Section II.1.

PROOF. If $\|g\|_2 = 0$, then Lemma 1.62 shows that $(f, g)_2 = 0$ for all f. Thus the lemma is valid in this case. If $\|g\|_2 \ne 0$, then we have

$$\begin{aligned}0 \le \big\|f - \|g\|_2^{-2}(f, g)_2\, g\big\|_2^2 &= \big(f - \|g\|_2^{-2}(f, g)_2\, g,\ f - \|g\|_2^{-2}(f, g)_2\, g\big)_2 \\ &= \|f\|_2^2 - 2\|g\|_2^{-2}|(f, g)_2|^2 + \|g\|_2^{-4}|(f, g)_2|^2\,\|g\|_2^2 = \|f\|_2^2 - \|g\|_2^{-2}|(f, g)_2|^2,\end{aligned}$$

and the lemma follows in this case as well.

Lemma 1.64 (triangle inequality). If f, g, and h are in $\mathcal{R}[-\pi, \pi]$, then $d_2(f, h) \le d_2(f, g) + d_2(g, h)$.

PROOF. For any two such functions F and G, Lemma 1.63 gives

$$\begin{aligned}\|F+G\|_2^2 &= (F+G, F+G)_2 = (F, F)_2 + (F, G)_2 + (G, F)_2 + (G, G)_2 \\ &= \|F\|_2^2 + 2\,\mathrm{Re}(F, G)_2 + \|G\|_2^2 \\ &\le \|F\|_2^2 + 2\|F\|_2\|G\|_2 + \|G\|_2^2 = (\|F\|_2 + \|G\|_2)^2.\end{aligned}$$

Taking the square root of both sides and substituting $F = f - g$ and $G = g - h$, we obtain the lemma.

Lemma 1.65. Let $f : \mathbb{R} \to \mathbb{C}$ be periodic of period 2π and Riemann integrable on $[-\pi, \pi]$, and let $\epsilon > 0$ be given. Then there exists a continuous periodic $g : \mathbb{R} \to \mathbb{C}$ of period 2π such that $\|f - g\|_2 < \epsilon$.

PROOF. Because of Lemma 1.64, we may assume that f is real-valued and is not identically 0. Define $M = \sup_{t\in[-\pi,\pi]} |f(t)| > 0$, let $\epsilon > 0$ be given, and let $P = \{x_i\}_{i=0}^n$ be a partition to be specified. Using P, we form the function g defined by

$$g(t) = \frac{x_i - t}{\Delta x_i}\, f(x_{i-1}) + \frac{t - x_{i-1}}{\Delta x_i}\, f(x_i) \quad \text{for } x_{i-1} \le t \le x_i.$$

The graph of g interpolates the points $(x_i, f(x_i))$, $0 \le i \le n$, by line segments. Fix attention on a particular $[x_{i-1}, x_i]$, and let $I = \inf_{t\in[x_{i-1},x_i]} f(t)$ and $S = \sup_{t\in[x_{i-1},x_i]} f(t)$. For $t \in [x_{i-1}, x_i]$, we have $I \le g(t) \le S$. At a single point t in this interval, $f(t) \ge g(t)$ implies $I \le g(t) \le f(t) \le S$, while $g(t) \ge f(t)$ implies $I \le g(t) \le f(t) \le S$. Thus in either case we have $|f(t) - g(t)| \le S - I$. Taking the supremum over t in the interval and summing on i, we obtain $U(P, |f - g|) \le U(P, f) - L(P, f)$.

Since $|f - g|^2 = |f - g||f + g|$, we have

$$\begin{aligned}\sup_{t\in[x_{i-1},x_i]} |f(t) - g(t)|^2 &\le \sup_{t\in[x_{i-1},x_i]} |f(t) - g(t)| \sup_{t\in[x_{i-1},x_i]} |f(t) + g(t)| \\ &\le 2M \sup_{t\in[x_{i-1},x_i]} |f(t) - g(t)|\end{aligned}$$

for $1 \le i \le n$. Summing on i gives $U\big(P, |f - g|^2\big) \le 2M(U(P, f) - L(P, f))$.

Now we can specify P; it is to be any partition for which $U(P, f) - L(P, f) \le \epsilon^2/(2M)$ and no Δx_i is 0. Then

$$\begin{aligned}0 &\le \tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} |f(t) - g(t)|^2\, dt \le \tfrac{1}{2\pi}\, U\big(P, |f - g|^2\big) \\ &\le \tfrac{2M}{2\pi}\, (U(P, f) - L(P, f)) \le \epsilon^2/(2\pi) < \epsilon^2,\end{aligned}$$

as required.

PROOF OF THEOREM 1.61. Given $\epsilon > 0$, choose by Lemma 1.65 a continuous periodic g with $\|f - g\|_2 < \epsilon$. Write $g(x) \sim \sum_{n=-\infty}^{\infty} c'_n e^{inx}$, and put $g_N(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(x-t)g(t)\,dt$, where K_N is the Fejér kernel. Fejér's Theorem (Theorem 1.59) gives $\sup_{x\in[-\pi,\pi]} |g(x)-g_N(x)| < \epsilon$ for N sufficiently large. Since any Riemann integrable h has $\|h\|_2 \le \sup_{x\in[-\pi,\pi]} |h(x)|$, we obtain $\|g - g_N\|_2 < \epsilon$ for N sufficiently large. Fixing such an N and substituting from the definition of K_N, we have

$$\begin{aligned} g_N(x) &= \frac{1}{N+1}\sum_{n=0}^{N}\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(x-t)g(t)\,dt \\ &= \frac{1}{N+1}\sum_{n=0}^{N}\sum_{k=-n}^{n} c'_k e^{ikx} = \sum_{n=-N}^{N} d_n e^{inx} \end{aligned}$$

for suitable constants d_n. Theorem 1.53 and Lemma 1.64 then give

$$\begin{aligned} \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx - \sum_{n=-N}^{N} |c_n|^2\Big)^{1/2} &= \Big\| f - \sum_{n=-N}^{N} c_n e^{inx}\Big\|_2 \\ &\le \Big\| f - \sum_{n=-N}^{N} d_n e^{inx}\Big\|_2 = \|f - g_N\|_2 \\ &\le \|f-g\|_2 + \|g - g_N\|_2 < \epsilon + \epsilon = 2\epsilon, \end{aligned}$$

and the result follows.

Corollary 1.66 (uniqueness theorem). Let $f : \mathbb{R} \to \mathbb{C}$ be periodic of period 2π and Riemann integrable on $[-\pi, \pi]$. If f has all Fourier coefficients 0, then $\int_{-\pi}^{\pi} |f(x)|\,dx = 0$ and $\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx = 0$ for every member g of $\mathcal{R}[-\pi, \pi]$.

PROOF. If f has all Fourier coefficients 0, then $\int_{-\pi}^{\pi} |f(x)|^2\,dx = 0$ by Theorem 1.61. Application of Lemma 1.62 completes the proof of the corollary.

It is natural to ask which sequences $\{c_n\}$ with $\sum |c_n|^2$ finite are the sequences of Fourier coefficients of some $f \in \mathcal{R}[-\pi, \pi]$. To see that this is a difficult question, one has only to compare the two series $\sum_{n=1}^{\infty} n^{-1}\sin x$ and $\sum_{n=1}^{\infty} n^{-1}\cos x$ studied at the beginning of this section. The first series comes from a function in $\mathcal{R}[-\pi, \pi]$, but a little argument shows that the second does not. It was an early triumph of Lebesgue integration that this question has a elegant answer when the Riemann integral is replaced by the Lebesgue integral: the answer when the Lebesgue integral is used is given by the Riesz–Fischer Theorem in Chapter VI, namely, *any* sequence with $\sum |c_n|^2$ finite is the sequence of Fourier coefficients of a square-integrable function.

11. Problems

1. Derive the least-upper-bound property (Theorem 1.1) from the convergence of bounded monotone increasing sequences (Corollary 1.6).

2. According to **Newton's method**, to find numerical approximations to $\sqrt{a}$ when $a > 0$, one can set $x_0 = 1$ and define $x_{n+1} = \frac{1}{2}(x_n^2 + a)/x_n$ for $n \geq 0$. Prove that $\{x_n\}$ converges and that the limit is $\sqrt{a}$.

3. Find $\limsup a_n$ and $\liminf a_n$ when a_n is defined by $a_1 = 0$, $a_{2n} = \frac{1}{2}a_{2n-1}$, $a_{2n+1} = \frac{1}{2} + a_{2n}$. Prove that your answers are correct.

4. For any two sequences $\{a_n\}$ and $\{b_n\}$ in $\mathbb{R}$, prove that $\limsup(a_n + b_n) \leq \limsup a_n + \limsup b_n$, provided the two terms on the right side are not $+\infty$ and $-\infty$ in some order.

5. Which of the following limits exist uniformly for $0 \leq x \leq 1$: (i) $\lim_{n\to\infty} x^n$, (ii) $\lim_{n\to\infty} x^n/n$, (iii) $\lim_{n\to\infty} \sum_{k=1}^n x^k/k$? Supply proofs for those that do converge uniformly. For the other ones, prove anyway that there is uniform convergence on any interval $0 \leq x \leq 1 - \epsilon$, where $\epsilon > 0$.

6. Let $a_n(x) = (-1)^n x^n(1 - x)$ on $[0, 1]$. Show that $\sum_{n=0}^\infty a_n(x)$ converges uniformly and that $\sum_{n=0}^\infty |a_n(x)|$ converges pointwise but not uniformly.

7. **(Dini's Theorem)** Suppose that $f_n : [a, b] \to \mathbb{R}$ is continuous and that $f_1 \leq f_2 \leq f_3 \leq \cdots$. Suppose also that $f(x) = \lim f_n(x)$ is continuous and is nowhere $+\infty$. Use the Bolzano–Weierstrass Theorem (Theorem 1.8) to prove that f_n converges to f uniformly for $a \leq x \leq b$.

8. Prove that
$$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \frac{x^{13}}{13!} - \frac{x^{15}}{15!} < \sin x$$
for all $x > 0$.

9. Let $f : (-\infty, +\infty) \to \mathbb{R}$ be infinitely differentiable with $|f^{(n)}(x)| \leq 1$ for all n and x. Use Taylor's Theorem (Theorem 1.36) to prove that
$$f(x) = \sum_{n=0}^\infty \frac{f^{(n)}(0)}{n!} x^n$$
for all x.

10. **(Helly's Selection Principle)** Suppose that $\{F_n\}$ is a sequence of nondecreasing functions on $[-1, 1]$ with $0 \leq F_n(x) \leq 1$ for all n and x. Using a diagonal process twice, prove that there is a subsequence $\{F_{n_k}\}$ that converges pointwise on $[-1, 1]$.

11. Prove that the radius of convergence of $\sum_{n=0}^\infty a_n x^n$ is $1/\limsup \sqrt[n]{|a_n|}$.

12. Find a power series expansion for each of the following functions, and find the radius of convergence:
    (a) $1/(1-x)^2 = \frac{d}{dx}(1-x)^{-1}$,
    (b) $\log(1-x) = -\int_1^x \frac{dt}{1-t}$,
    (c) $1/(1+x^2)$,
    (d) $\arctan x = \int_0^x \frac{dt}{1+t^2}$.

13. Prove, along the lines of the proof of Corollary 1.46a, that $\cos x$ has an inverse function $\arccos x$ defined for $0 < x < \pi$ and that the inverse function is differentiable. Find an explicit formula for the derivative of $\arccos x$. Relate $\arccos x$ to $\arcsin x$ when $0 < x < \pi/2$.

14. State and prove uniform versions of Abel's Theorem (Theorem 1.48) and of the corresponding theorem about Cesàro sums (Theorem 1.47), the uniformity being with respect to a parameter x.

15. Prove that the partial sums $\sum_{n=1}^N \cos n\theta$ and $\sum_{n=1}^N \sin n\theta$ are uniformly bounded on any set $\epsilon \le \theta < 2\pi - \epsilon$ if $\epsilon > 0$.

16. Verify the following calculations of Fourier series:
    (a) $f(x) = \left\{ \begin{matrix} +1 & \text{for } 0 < x < \pi \\ -1 & \text{for } -\pi < x < 0 \end{matrix} \right\}$ has $f(x) \sim \dfrac{4}{\pi} \displaystyle\sum_{n=1}^\infty \dfrac{\sin(2n-1)x}{2n-1}$.
    (b) $f(x) = e^{-i\alpha x}$ on $(0, 2\pi)$ has $f(x) \sim \dfrac{e^{-i\pi\alpha} \sin \pi\alpha}{\pi} \displaystyle\sum_{n=-\infty}^\infty \dfrac{e^{inx}}{n+\alpha}$, provided α is not an integer.

17. Combining Parseval's Theorem (Theorem 1.61) with the results of Problem 16, prove the following identities:
    (a) $\displaystyle\sum_{n=1}^\infty \frac{1}{(2n-1)^2} = \frac{\pi^2}{8}$, (b) $\displaystyle\sum_{n=-\infty}^\infty \frac{1}{|n+\alpha|^2} = \frac{\pi^2}{\sin^2 \pi\alpha}$.

Problems 18–19 identify the continuous functions $f : \mathbb{R} \to \mathbb{C}$ with $f(x)f(y) = f(x+y)$ for all x and y as the 0 function and the functions $f(x) = e^{cx}$, using two different kinds of techniques from the chapter.

18. Put $F(x) = \int_0^x f(t)\,dt$. Find an equation satisfied by F, and use it to show that f is differentiable everywhere. Then show that $f'(y) = f'(0)f(y)$, and deduce the form of f.

19. Proceed without using integration. Using continuity, find $x_0 > 0$ such that the expression $|f(x) - 1|$ is suitably small when $|x| \le |x_0|$. Show that $f(2^{-k}x_0)$ is then uniquely determined in terms of $f(x_0)$ for all $k \ge 0$. If f is not identically 0, use x_0 to define c. Then verify that $f(x) = e^{cx}$ for all c.

Problems 20–22 construct a nonzero infinitely differentiable function $f : \mathbb{R} \to \mathbb{R}$ having all derivatives equal to 0 at one point.

20. Let $P(x)$ and $Q(x)$ be two polynomials with Q not the zero polynomial. Prove that
$$\lim_{x\to 0} \tfrac{P(x)}{Q(x)}\, e^{-1/x^2} = 0.$$

21. With P and Q as in the previous problem, use the Mean Value Theorem to prove that the function $g : \mathbb{R} \to \mathbb{R}$ with
$$g(x) = \begin{cases} \tfrac{P(x)}{Q(x)}\, e^{-1/x^2} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0, \end{cases}$$
has $g'(x) = 0$ and that g' is continuous.

22. Prove that the function $f : \mathbb{R} \to \mathbb{R}$ with
$$f(x) = \begin{cases} e^{-1/x^2} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0, \end{cases}$$
is infinitely differentiable with derivatives of all orders equal to 0 at $x = 0$.

Problems 23–26 concern a generalization of Cesàro and Abel summability. A **Silverman–Toeplitz summability method** refers to the following construction: One starts with a system $\{M_{ij}\}_{i,j\ge 0}$ of nonnegative real numbers with the two properties that (i) $\sum_j M_{ij} = 1$ for all i and (ii) $\lim_{i\to\infty} M_{ij} = 0$ for all j. The method associates to a complex sequence $\{s_n\}_{n\ge 0}$ the complex sequence $\{t_n\}_{n\ge 0}$ with $t_i = \sum_{j\ge 0} M_{ij} s_j$ as if the process were multiplication by the infinite square matrix $\{M_{ij}\}$ on infinite column vectors.

23. Prove that if $\{s_n\}$ is a convergent sequence with limit s, then the corresponding sequence $\{t_n\}$ produced by a Silverman–Toeplitz summability method converges and has limit s.

24. Exhibit specific matrices $\{M_{ij}\}$ that produce the effects of Cesàro and Abel summability, the latter along a sequence r_i increasing to 1.

25. Let r_i be a sequence increasing to 1, and define $M_{ij} = (j+1)(r_i)^j(1-r_i)^2$. Show that $\{M_{ij}\}$ defines a Silverman–Toeplitz summability method.

26. Using the system $\{M_{ij}\}$ in the previous problem, prove the following: if a bounded sequence $\{s_n\}$ is not necessarily convergent but is Cesàro summable to a limit σ, then $\{s_n\}$ is Abel summable to the same limit σ.

Problems 27–29 concern the Poisson kernel, which plays the same role for Abel sums of Fourier series that the Fejér kernel plays for Cesàro sums. For $0 \le r < 1$, define the **Poisson kernel** $P_r(\theta)$ to be the r^{th} Abel sum of the Dirichlet kernel $D_n(\theta) = 1 + \sum_{k=1}^n (e^{ik\theta} + e^{-ik\theta})$. In the terminology of Section 8 this means that $a_0 = 1$ and $a_k = e^{ik\theta} + e^{-ik\theta}$ for $k \ge 0$, so that the sequence of partial sums $\sum_{k=0}^n a_k$ is exactly the sequence whose n^{th} term is $D_n(\theta)$. The r^{th} Abel sum $\sum_{n=0}^{\infty} a_n r^n$ is therefore the expression

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta}.$$

27. For f in $\mathcal{R}[-\pi, \pi]$, verify that the r^{th} Abel sum of $s_n(f; x)$ is given by the expression $\frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(\theta - \varphi) f(\varphi)\, d\varphi$.

28. Verify that $P_r(\theta) = \dfrac{1 - r^2}{1 - 2r\cos\theta + r^2}$. Deduce that $P_r(\theta)$ has the following properties:
    (i) $P_r(\theta) \geq 0$,
    (ii) $\frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(\theta)\, d\theta = 1$,
    (iii) for any $\delta > 0$, $\sup_{\delta \leq |\theta| \leq \pi} P_r(\theta)$ tends to 0 as r increases to 1.

29. Let $f : \mathbb{R} \to \mathbb{C}$ be periodic of period 2π and Riemann integrable on $[-\pi, \pi]$.
    (a) Prove that if f is continuous at a point θ_0 in $[-\pi, \pi]$, then

$$\lim_{r \uparrow 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(\theta_0 - \theta) f(\theta)\, d\theta = f(\theta_0).$$

    (b) Prove that if f is uniformly continuous on a subset E of $[-\pi, \pi]$, then the convergence in (a) is uniform for θ_0 in E.

Problems 30–35 lead to a proof without complex-variable theory (and in particular without the complex logarithm) that $\exp\big(z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \cdots\big) = 1/(1 - z)$ for all complex z with $|z| < 1$.

30. Suppose that $R > 0$, that $f_k(x) = \sum_{n=0}^{\infty} c_{n,k} x^n$ is convergent for $|x| < R$, that $c_{n,k} \geq 0$ for all n and k, and that $\lim_{k\to\infty} f_k(x) = f(x)$ uniformly for $|x| \leq r$ whenever $r < R$. Prove for each $r < R$ that some subsequence $\{f_{k_l}\}$ of $\{f_k\}$ has $\lim_{l\to\infty} f'_{k_l}(x)$ existing uniformly for $|x| \leq r$.

31. In the setting of the previous problem, prove that f is infinitely differentiable for $|x| < R$.

32. In the setting of the previous two problems, use Taylor's Theorem to show that $f(x)$ is the sum of its infinite Taylor series for $|x| < R$.

33. If $0 \leq r < 1$, prove for $|z| \leq r$ that $\big|\frac{1}{N} z^N + \frac{1}{N+1} z^{N+1} + \cdots\big| \leq r^N/(1 - r)$, and deduce that $\exp\big(\frac{1}{N} z^N + \frac{1}{N+1} z^{N+1} + \cdots\big)$ converges to 1 uniformly for $|z| \leq r$.

34. Why is it true that if a power series $\sum_{n=0}^{\infty} c_n z^n$ with complex coefficients sums to 0 for all real z with $|z| < R$, then it sums to 0 for all complex z with $|z| < R$?

35. Prove that $\exp\big(z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \cdots\big) = 1/(1 - z)$ for all complex z with $|z| < 1$.

CHAPTER II

Metric Spaces

Abstract. This chapter is about metric spaces, an abstract generalization of the real line that allows discussion of open and closed sets, limits, convergence, continuity, and similar properties. The usual distance function for the real line becomes an example of a metric. The other notions are defined in terms of the metric. The advantage of the generalization is that proofs of certain properties of the real line immediately go over to all other examples.

Section 1 gives the definition of metric space and open set, and it lists a number of important examples, including Euclidean spaces and certain spaces of functions.

Sections 2 through 4 develop properties of open and closed sets, continuity, and convergence of sequences that are simple generalizations of known facts about $\mathbb{R}$.

Section 5 shows how a subset of a metric space can be made into a metric space so that the restriction of a continuous function from the whole space to the subset remains continuous. It also shows that three natural metrics for the product of two metric spaces lead to the same open sets, continuous functions, and convergent sequences.

Section 6 shows that any metric space is "Hausdorff," "regular," and "normal," and it goes on to exhibit three different countability hypotheses about a metric space as equivalent. A metric space with these properties is called "separable."

Section 7 concerns compactness and completeness. A metric space is defined to be "compact" if every open cover has a finite subcover. This property is equivalent to the condition that every sequence has a convergent subsequence. The Heine–Borel Theorem says that the compact sets of $\mathbb{R}^n$ are exactly the closed bounded sets. A number of the results early in Chapter I that were proved by the Bolzano–Weierstrass Theorem in the context of the real line are seen to extend to any compact metric space. A metric space is "complete" if every Cauchy sequence is convergent. A metric space is compact if and only if it is complete and "totally bounded."

Section 8 concerns connectedness, which is an abstraction of the property of an interval of the line that accounts for the Intermediate Value Theorem.

Section 9 proves a fundamental result known as the Baire Category Theorem. A sample consequence of the theorem is that the pointwise limit of a sequence of continuous complex-valued functions on a complete metric space must have points where it is continuous.

Section 10 studies the spaces of real-valued and complex-valued continuous functions on a compact metric space. A generalization of Ascoli's Theorem from the setting of Chapter I provides a characterization of compact sets in either of these spaces of continuous functions. A generalization of the Weierstrass Approximation Theorem, known as the Stone–Weierstrass Theorem, gives sufficient conditions for a subalgebra of either of these spaces of continuous functions to be dense. One consequence is that these spaces of continuous functions are separable.

Section 11 constructs the "completion" of a metric space out of Cauchy sequences in the given space. The result is a complete metric space and a distance-preserving map of the given metric space into the completion such that the image is dense.

1. Definition and Examples

Let X be a nonempty set. A function d from $X \times X$, the set of ordered pairs of members of X, to the real numbers is a **metric**, or distance function, if

(i) $d(x, y) \geq 0$ always, with equality if and only if $x = y$,
(ii) $d(x, y) = d(y, x)$ for all x and y in X,
(iii) $d(x, y) \leq d(x, z) + d(z, y)$ for all x, y, and z, the **triangle inequality**.

In this case the pair (X, d) is called a **metric space**.

The real line $\mathbb{R}^1$ with metric $d(x, y) = |x - y|$ is the motivating example. Properties (i) and (ii) are apparent, and property (iii) is readily verified one case at a time according as z is less than both x and y, z is between x and y, or z is greater than both x and y.

We come to further examples in a moment. Particularly in the case that X is a space of functions, a space may turn out to be almost a metric space but not to satisfy the condition that $d(x, y) = 0$ implies $x = y$. Accordingly we introduce a weakened version of (i) as

(i′) $d(x, y) \geq 0$ and $d(x, x) = 0$ always,

and we say that a function d from $X \times X$ to the real numbers is a **pseudometric** if (i′), (ii), and (iii) hold. In this case, (X, d) is called a **pseudometric space**.

Let (X, d) be a pseudometric space. If $r > 0$, the **open ball** of radius r and center x, denoted by $B(r; x)$, is the set of points at distance less than r from x, namely

$$B(r; x) = \{y \in X \mid d(x, y) < r\}.$$

The name "ball" will be appropriate in Euclidean space in dimension three, which is part of the Example 1 below, and "ball" is adopted for the corresponding notion in a general pseudometric space.

A subset U of X is **open** if for each x in U and some sufficiently small $r > 0$, the open ball $B(r; x)$ is contained in U. For the line the open balls in the above sense are just the bounded open intervals, and the open sets in the above sense are the usual open sets in the sense of Chapter I.

Lemma 2.1. In any pseudometric space (X, d), every open ball is an open set. The open sets are exactly all possible unions of open balls.

PROOF. Let an open ball $B(r; x)$ be given. If y is in $B(r; x)$, then the open ball $B(r - d(x, y),\ y)$ has center y and positive radius; we show that it is contained in $B(r; x)$. In fact, if z is in $B(r - d(x, y),\ y)$, then the triangle inequality gives

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + (r - d(x, y)) = r,$$

and the containment follows.

For the second assertion it follows from the definition of open set that every open set is the union of open balls. In the reverse direction, let U be a union of open balls. If y is in U, then y lies in one of these balls, say in $B(r; x)$. We have just shown that some open ball $B(s; y)$ is contained in $B(r; x)$, and $B(r; x)$ is contained in U. Thus $B(s; y)$ is contained in U, and U is open.

EXAMPLES.

(1) **Euclidean space** $\mathbb{R}^n$. Fix an integer $n > 0$. Let $\mathbb{R}^n$ be the space of all n-tuples of real numbers $x = (x_1, \dots, x_n)$. We define addition of n-tuples componentwise, and we define scalar multiplication by $cx = (cx_1, \dots, cx_n)$ for real c. Following the normal convention in linear algebra, we identify this space with the real vector space, also denoted by $\mathbb{R}^n$, of all n-component column vectors of real numbers $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Generalizing the notion of absolute value when $n = 1$, we let $|x| = \big(\sum_{j=1}^n x_j^2\big)^{1/2}$ for $x = (x_1, \dots, x_n)$ in $\mathbb{R}^n$. The quantity $|x|$ is the **Euclidean norm** of x. The Euclidean norm satisfies the properties

(a) $|x| \geq 0$ always, with equality if and only if x equals the zero tuple $0 = (0, \dots, 0)$,
(b) $|cx| = |c||x|$ for all x and for all real c,
(c) $|x + y| \leq |x| + |y|$ for all x and y.

Properties (a) and (b) are apparent, but (c) requires proof. The proof makes use of the familiar **dot product**, given by $x \cdot y = \sum_{j=1}^n x_j y_j$ if $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$. In terms of dot product, the Euclidean norm is nothing more than $|x| = (x \cdot x)^{1/2}$. The dot product satisfies the important inequality $|x \cdot y| \leq |x||y|$, known as the **Schwarz inequality** and proved for this context in Section A5 of the appendix. A more general version of the Schwarz inequality will be stated and proved in Lemma 2.2 below. The Schwarz inequality implies (c) above because we then have

$$\begin{aligned} |x + y|^2 &= (x + y) \cdot (x + y) = x \cdot x + 2(x \cdot y) + y \cdot y \\ &= |x|^2 + 2(x \cdot y) + |y|^2 \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2. \end{aligned}$$

We make $X = \mathbb{R}^n$ into a metric space (X, d) by defining

$$d(x, y) = |x - y|.$$

Properties (i) and (ii) of a metric are immediate from (a) and (b), respectively; property (iii) follows from (c) in the form $|a + b| \leq |a| + |b|$ if we substitute $a = x - z$ and $b = y - z$. For $n = 1$, this example reduces to the line as discussed above. For $n = 2$, open balls are geometric open disks, while for $n = 3$, open balls are geometric open balls. For any n, the open sets in the metric space coincide with the open sets as defined in calculus of several variables.

(2) **Complex Euclidean space** $\mathbb{C}^n$. The space $\mathbb{C}$ of complex numbers, with distance function $d(z, w) = |z - w|$ as in Section I.5, can be seen in two ways to be a metric space. One way was carried out in Section I.5 and directly uses the properties of the absolute value function $|z|$ in Section A4 of the appendix. The other way is to identify $z = x + iy$ with the member (x, y) of $\mathbb{R}^2$, and then the absolute value $|z|$ equals the Euclidean norm $|(x, y)|$ in the sense of Example 1; hence the construction of Example 1 makes the set of complex numbers into a metric space. More generally the complex vector space $\mathbb{C}^n$ of n-tuples

$$z = (z_1, \dots, z_n) = (x_1, \dots, x_n) + i(y_1, \dots, y_n) = x + iy$$

becomes a metric space in two equivalent ways. One way is to define the norm $|z| = \big(\sum_{j=1}^n |z_j|^2\big)^{1/2}$ as a generalization of the Euclidean norm for $\mathbb{R}^n$; then we put $d(z, w) = |z - w|$. The argument that d satisfies the triangle inequality is a variant of the one for $\mathbb{R}^n$: The object for $\mathbb{C}^n$ that generalizes the dot product for $\mathbb{R}^n$ is the **Hermitian inner product**

$$(z, w) = \big((z_1, \dots, z_n), (w_1, \dots, w_n)\big) = \sum_{j=1}^n z_j \overline{w_j}.$$

The **Euclidean norm** is given in terms of this expression by $|z| = (z, z)^{1/2}$, and the version of the Schwarz inequality in Section A5 of the appendix is general enough to show that $|(z, w)| \le |z||w|$. The same argument as for Example 1 shows that the norm satisfies the triangle inequality, and then it follows that d satisfies the triangle inequality. The other way to view $\mathbb{C}^n$ as a metric space is to identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ by $(z_1, \dots, z_n) \mapsto (x_1, \dots, x_n, y_1, \dots, y_n)$ and then to use the metric on $\mathbb{R}^{2n}$ from Example 1. This is the same metric, since $\sum_{j=1}^n |z_j|^2 = \sum_{j=1}^n x_j^2 + \sum_{j=1}^n y_j^2$. We still get the same metric if we instead use the identification $(z_1, \dots, z_n) \mapsto (x_1, y_1, \dots, x_n, y_n)$. With either identification the Hermitian inner product (z, w) for $\mathbb{C}^n$ corresponds to the ordinary dot product for $\mathbb{R}^{2n}$.

(3) System $\mathbb{R}^*$ of extended real numbers. The function $f(x) = x/(1 + x)$ carries $[0, +\infty)$ into $[0, +1)$ and has $g(y) = y/(1 - y)$ as a two-sided inverse. Therefore f is one-one and onto. We can extend f so that it carries $(-\infty, +\infty)$ one-one onto $(-1, +1)$ by putting $f(x) = x/(1 + |x|)$. We can extend f further by putting $f(-\infty) = -1$ and $f(+\infty) = +1$, and then f carries $[-\infty, +\infty]$, i.e., all of $\mathbb{R}^*$, one-one onto $[-1, +1]$. The function f is nondecreasing on $[-\infty, +\infty]$. For x and x' in $\mathbb{R}^*$, let

$$d(x, x') = |f(x) - f(x')|.$$

We shall show that d is a metric. By inspection, d satisfies properties (i) and (ii) of a metric, and we are to prove the triangle inequality (iii), namely that

$$d(x, x') \le d(x, x'') + d(x'', x').$$

The critical fact is that f is nondecreasing. Since d satisfies (ii), we may assume that $x \le x'$, and then

$$d(x, x') = f(x') - f(x).$$

We divide the proof into three cases, depending on the location of x'' relative to x and x'. The first case is that $x'' \le x$, and then

$$d(x, x'') + d(x'', x') = f(x) - f(x'') + f(x') - f(x'').$$

Thus the question is whether

$$f(x') - f(x) \stackrel{?}{\le} f(x) - f(x'') + f(x') - f(x''),$$

hence whether

$$2f(x'') \stackrel{?}{\le} 2f(x).$$

This inequality holds, since f is nondecreasing. The second case is that $x \le x'' \le x'$, and then

$$d(x, x'') + d(x'', x') = f(x'') - f(x) + f(x') - f(x'') = f(x') - f(x).$$

Hence equality holds in the triangle inequality. The third case is that $x' \le x''$, and then

$$d(x, x'') + d(x'', x') = f(x'') - f(x) + f(x'') - f(x').$$

The triangle inequality comes down to the question whether

$$2f(x') \stackrel{?}{\le} 2f(x'').$$

This inequality holds, since f is nondecreasing. We conclude that $(\mathbb{R}^*, d)$ is a metric space. It is not hard to see that the open balls in $\mathbb{R}^*$ are all intervals (a, b), $[-\infty, b)$, $(a, +\infty]$, and $[-\infty, +\infty]$ with $-\infty \le a < b \le +\infty$. Each of these open balls in $\mathbb{R}^*$ intersects $\mathbb{R}$ in an ordinary open interval, bounded or unbounded. The open sets in $\mathbb{R}$ therefore coincide with the intersections of $\mathbb{R}$ with the open sets of $\mathbb{R}^*$.

(4) Bounded functions in the **uniform metric**. Let S be a nonempty set, and let $X = B(S)$ be the set of all "scalar"-valued functions f on S that are **bounded** in the sense that $|f(s)| \le M$ for all $s \in S$ and for a constant M depending on f. The **scalars** are allowed to be the members of $\mathbb{R}$ or the members of $\mathbb{C}$, and it will ordinarily make no difference which one is understood. If it does make a difference, we shall write $B(S, \mathbb{R})$ or $B(S, \mathbb{C})$ to be explicit about the range. For f and g in $B(S)$, let

$$d(f, g) = \sup_{s \in S} |f(s) - g(s)|.$$

It is easy to verify that (X, d) is a metric space. Let us not lose sight of the fact that the members of X are functions. When we discuss convergence of sequences in a metric space, we shall see that a sequence of functions in this X converges if and only if the sequence of functions converges uniformly on S.

(5) Generalization of Example 4. We can replace the range $\mathbb{R}$ or $\mathbb{C}$ of the functions in Example 4 by any metric space (R, ρ). Fix a point r_0 in the range R. A function $f : S \to R$ is **bounded** if $\rho(f(s), r_0) \le M$ for all s and for some M depending on f. This definition is independent of the choice of r_0 because ρ is assumed to satisfy the triangle inequality. If we let X be the space of all such bounded functions from S to R, we can make X into a metric space by defining $d(f, g) = \sup_{s \in S} \rho(f(s), g(s))$.

(6) Sequence space ℓ^2. This is the space of all sequences $\{c_n\}_{n=-\infty}^{\infty}$ of scalars with $\sum |c_n|^2 < \infty$. A metric is given by

$$d(\{c_n\}, \{d_n\}) = \Big(\sum_{n=-\infty}^{\infty} |c_n - d_n|^2 \Big)^{1/2}.$$

In the case of complex scalars, this example arises as a natural space containing all systems of Fourier coefficients of Riemann integrable functions on $[-\pi, \pi]$, in the sense of Chapter I. Proving the triangle inequality involves arguing as in Examples 1 and 2 above and then letting the number of terms tend to infinity. The role of the dot product is played by $(\{c_n\}, \{d_n\}) = \sum_{n=-\infty}^{\infty} c_n \overline{d_n}$.

(7) **Indiscrete space.** If X is any nonempty set and if $d(x, y) = 0$ for all x and y, then d is a pseudometric and the only open sets are X and the empty set $\varnothing$. If X contains more than one element, then d is not a metric.

(8) **Discrete metric.** If X is any nonempty set and if

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y, \\ 0 & \text{if } x = y, \end{cases}$$

then d is a metric, and every subset of X is open.

(9) Let S be a nonempty set, fix an integer $n > 0$, and let X be the set of n-tuples of members of S. For n-tuples $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, define

$$d(x, y) = \#\{j \mid x_j \neq y_j\},$$

the number of components in which x and y differ. Then (X, d) is a metric space. The proof of the triangle inequality requires a little argument, but we leave that for Problem 1 at the end of the chapter. Every subset of X is open, just as with the discrete metric in Example 8.

(10) **Hedgehog space.** Let X be $\mathbb{R}^2$, and single out the origin for special attention. Let d be the metric of Euclidean space, and define

$$\rho(x, y) = \begin{cases} d(x, y) & \text{if } x \text{ and } y \text{ are on the same ray from } 0, \\ d(x, 0) + d(0, y) & \text{otherwise.} \end{cases}$$

Then ρ is a metric. Every open set in (X, d) is open in (X, ρ), but a set like the one in Figure 2.1 is open in (X, ρ) but not in (X, d).

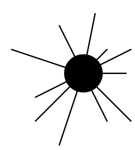


FIGURE 2.1. An open set centered at the origin in the hedgehog space.

(11) **Hilbert cube.** Let X be the set of all sequences $\{x_m\}_{m\geq 1}$ of real numbers satisfying $0 \leq x_m \leq 1$ for all m, and put

$$d(\{x_m\}, \{y_m\}) = \sum_{m=1}^{\infty} 2^{-m} |x_m - y_m|.$$

Then (X, d) is a metric space. To verify the triangle inequality, we can argue as follows: Let $\{x_m\}$, $\{y_m\}$, and $\{z_m\}$ be in X. For each m, we have

$$2^{-m}|x_m - y_m| \leq 2^{-m}|x_m - z_m| + 2^{-m}|z_m - y_m|.$$

Thus

$$\begin{aligned} \sum_{m=1}^{N} 2^{-m}|x_m - y_m| &\leq \sum_{m=1}^{N} 2^{-m}|x_m - z_m| + \sum_{m=1}^{N} 2^{-m}|z_m - y_m| \\ &\leq \sum_{m=1}^{\infty} 2^{-m}|x_m - z_m| + \sum_{m=1}^{\infty} 2^{-m}|z_m - y_m| \end{aligned}$$

for each N. Letting N tend to infinity yields the desired inequality.

(12) L^1 metric on Riemann integrable functions. Fix a nontrivial bounded interval $[a, b]$ of the line, let X be the set of all Riemann integrable complex-valued functions on $[a, b]$ in the sense of Chapter I, and define

$$d_1(f, g) = \int_a^b |f(x) - g(x)|\, dx$$

for f and g in X. Then (X, d_1) is a pseudometric space. It can happen that $\int_a^b |f(x) - g(x)|\, dx = 0$ without $f = g$; for example, f could differ from g at a single point. Therefore d_1 is not a metric.

(13) L^2 metric on complex-valued $\mathcal{R}[-\pi, \pi]$. This example arose in the discussion of Fourier series in Section I.10, and it was convenient to include a factor $\frac{1}{2\pi}$ in front of integrals. Let $X = \mathcal{R}[-\pi, \pi]$, and define

$$d_2(f, g) = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - g(x)|^2\, dx\Big)^{1/2}.$$

Then (X, d_2) is a pseudometric metric space. The triangle inequality was proved in Lemma 1.64 using the version of the Schwarz inequality in Lemma 1.63; that version of the Schwarz inequality needed a special argument given in Lemma 1.62 in order to handle functions f whose norm satisfies $\|f\|_2 = 0$.

The constructions of metric spaces in Examples 1, 2, 6, and 13 are sufficiently similar to warrant abstracting what was involved. We start with a real or complex vector space V, possibly infinite-dimensional, and with a generalization $(\,\cdot\,,\,\cdot\,)$ of dot product. This generalization is a function from $V \times V$ to $\mathbb{R}$ in the case that V is real, and it is a function from $V \times V$ to $\mathbb{C}$ in the case that V is complex. We shall write the scalars as if they are complex, but only real scalars are to be used if the vector space is real. The function is written $(\,\cdot\,,\,\cdot\,)$ and is assumed to satisfy the following properties:

(i) it is linear in the first variable, i.e., $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$ and $(cx, y) = c(x, y)$,

(ii) it is conjugate linear in the second variable, i.e., $(x, y_1 + y_2) = (x, y_1) + (x, y_2)$ and $(x, cy) = \bar{c}(x, y)$,

(iii) it is **symmetric** in the real case and **Hermitian symmetric** in the complex case, i.e., $(y, x) = \overline{(x, y)}$,

(iv) it is **definite**, i.e., $(x, x) > 0$ if $x \neq 0$.

The form $(\,\cdot\,,\,\cdot\,)$ is called an **inner product** if V is real or complex and is often called also a **Hermitian inner product** if V is complex; in either case, V with the form is called an **inner-product space**. Two vectors x and y with $(x, y) = 0$ are said to be **orthogonal**; the notion of orthogonality generalizes perpendicularity in the case of the dot product.

For either kind of scalars, we define $\|x\| = (x, x)^{1/2}$, and the function $\|\cdot\|$ is called the associated **norm**. We shall see shortly that a version of the Schwarz inequality is valid in this generality, the proof being no more complicated than the one in Section A5 of the appendix.

In many cases in practice, item (iv) is replaced by the weaker condition that

(iv′) $(\cdot\,,\,\cdot)$ is **semidefinite**, i.e., $(x, x) \geq 0$ if $x \neq 0$.

This was what happened in Example 13 above. In order to have a name for this kind of space, let us call V with the semidefinite form $(\cdot\,,\,\cdot)$ a **pseudo inner-product space**. It is still meaningful to speak of orthogonality. It is still meaningful also to define $\|x\| = (x, x)^{1/2}$, and this is called the **pseudonorm** for the space. The Schwarz inequality is still valid, but its proof is more complicated than for an inner-product space. The extra complication was handled by Lemma 1.62 in the case of Example 13 in order to obtain a little extra information; the general argument proceeds along different lines.

Lemma 2.2 (Schwarz inequality). Let V be a pseudo inner-product space with form $(\cdot\,,\,\cdot)$. If x and y are in V, then $|(x, y)| \leq \|x\|\|y\|$.

PROOF. First suppose that $\|y\| \neq 0$. Then

$$\begin{aligned} 0 &\leq \big|x - \|y\|^{-2}(x, y)y\big|^2 = \big((x - \|y\|^{-2}(x, y)y),\ (x - \|y\|^{-2}(x, y)y)\big) \\ &= \|x\|^2 - 2\|y\|^{-2}|(x, y)|^2 + \|y\|^{-4}|(x, y)|^2\|y\|^2 = \|x\|^2 - \|y\|^{-2}|(x, y)|^2, \end{aligned}$$

and the inequality follows in this case.

Next suppose that $\|y\| = 0$. It is enough to prove that $(x, y) = 0$ for all x. If c is a real scalar, we have

$$\|x+cy\|^2 = (x+cy, x+cy) = \|x\|^2 + 2\operatorname{Re}(x, cy) + |c|^2\|y\|^2 = \|x\|^2 + 2c\operatorname{Re}(x, y).$$

The left side is ≥ 0 as c varies, but the right side can be < 0 unless $\operatorname{Re}(x, y) = 0$. Thus we must have $\operatorname{Re}(x, y) = 0$ for all x. Replacing x by ix gives us $\operatorname{Im}(x, y) = -\operatorname{Re} i(x, y) = -\operatorname{Re}(ix, y)$, and this we have just shown is 0 for all x. Thus $\operatorname{Re}(x, y) = \operatorname{Im}(x, y) = 0$, and $(x, y) = 0$.

Proposition 2.3 (triangle inequality). If V is a pseudo inner-product space with form $(\cdot\,,\,\cdot)$ and pseudonorm $\|\cdot\|$, then the pseudonorm satisfies

(a) $\|x\| \geq 0$ for all $x \in V$,
(b) $\|cx\| = |c|\|x\|$ for all scalars c and all $x \in V$,
(c) $\|x + y\| \leq \|x\| + \|y\|$ for all x and y in V.

Moreover, the definition $d(x, y) = \|x - y\|$ makes V into a pseudometric space. The space V is a metric space if the pseudo inner-product space is an inner-product space.

PROOF. Properties (a) and (b) of the pseudonorm are immediate, and (c) follows because

$$\begin{aligned}\|x+y\|^2 &= (x+y,\ x+y) = (x,x) + 2\,\mathrm{Re}(x,y) + (y,y)\\ &= \|x\|^2 + 2\,\mathrm{Re}(x,y) + \|y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.\end{aligned}$$

Putting $x = a - c$ and $y = c - b$ gives $d(a,b) \le d(a,c) + d(c,b)$, and thus d satisfies the triangle inequality for a pseudometric. The other properties of a pseudometric are immediate from (a) and (b). If the form is definite and $d(f,g) = 0$, then $(f-g,\ f-g) = 0$ and hence the definiteness yields $f - g = 0$.

EXAMPLES, continued.

14) Let us take double integrals of continuous functions of nice subsets of $\mathbb{R}^2$ as known. (The detailed study of general Riemann integrals in several variables occurs in Chapter III.) Let V be the complex vector space of all power series $F(z) = \sum_{n=0}^\infty c_n z^n$ with infinite radius of convergence. Since any such $F(z)$ is bounded on the open unit disk $D = \{z \in \mathbb{C} \mid |z| < 1\}$, the form $(F,G) = \int_D F(z)\overline{G(z)}\,dx\,dy$ is meaningful and makes V into an inner-product space. The proposition shows that V becomes a metric space with metric given by $d(F,G) = \big(\int_D |F(z) - G(z)|^2\,dx\,dy\big)^{1/2}$.

2. Open Sets and Closed Sets

In this section we generalize the Euclidean notions of open set, closed set, neighborhood, interior, limit point, and closure so that they make sense for all pseudometric spaces, and we prove elementary properties relating these metric-space notions. In working with metric spaces and pseudometric spaces, it is often helpful to draw pictures as if the space in question were $\mathbb{R}^2$, even computing distances that are right for $\mathbb{R}^2$. We shall do that in the case of the first lemma but not afterward in this section. Let (X,d) be a pseudometric space.

Lemma 2.4. If z is in the intersection of open balls $B(r;x)$ and $B(s;y)$, then there exists some $t > 0$ such that the open ball $B(t;z)$ is contained in that intersection. Consequently the intersection of two open balls is open.

REMARK. Figure 2.2 shows what $B(t;z)$ looks like in the metric space $\mathbb{R}^2$.

PROOF. Take $t = \min\{r - d(x,z),\ s - d(y,z)\}$. If w is in $B(t;z)$, then the triangle inequality gives

$$d(x,w) \le d(x,z) + d(z,w) < d(x,z) + t \le d(x,z) + (r - d(x,z)) = r,$$

and hence w is in $B(r;x)$. Similarly w is in $B(s;y)$.

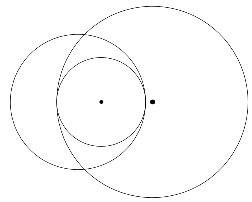


FIGURE 2.2. Open ball contained in an intersection of two open balls.

Proposition 2.5. The open sets of X have the properties that

(a) X and the empty set $\varnothing$ are open,
(b) an arbitrary union of open sets is open,
(c) any finite intersection of open sets is open.

PROOF. We know from Lemma 2.1 that a set is open if and only if it is the union of open balls. Then (b) is immediate, and (a) follows, since X is the union of all open balls and $\varnothing$ is an empty union. For (c), it is enough to prove that $U \cap V$ is open if U and V are open. Write $U = \bigcup_\alpha B_\alpha$ and $V = \bigcup_\beta B_\beta$ as unions of open balls. Then $U \cap V = \bigcup_{\alpha,\beta} (B_\alpha \cap B_\beta)$, and Lemma 2.4 shows that $U \cap V$ is exhibited as the union of open balls. Thus $U \cap V$ is open.

A **neighborhood** of a point in X is any set that contains an open set containing the point. An **open neighborhood** is a neighborhood that is an open set.[1] A **neighborhood** of a subset E of X is a set that is a neighborhood of each point of E. If A is a subset of X, then the set A^o of all points x in A for which A is a neighborhood of x is called the **interior** of A. For example, the interior of the half-open interval $[a, b)$ of the real line is the open interval (a, b).

Proposition 2.6. The interior of a subset A of X is the union of all open sets contained in A; that is, it is the largest open set contained in A.

PROOF. Suppose that $U \subseteq A$ is open. If x is in U, then U is an open neighborhood of x, and hence A is a neighborhood of x. Thus x is in A^o, and A^o contains the union of all open sets contained in A. For the reverse inclusion, let x be in A^o. Then A is a neighborhood of x, and there exists an open subset U of A containing x. So x is contained in the union of all open sets contained in A.

Corollary 2.7. A subset A of X is open if and only if $A = A^o$.

A subset F of X is **closed** if its complement is open. Every closed interval of the real line is closed. A half-open interval $[a, b)$ on the real line is neither open nor closed if a and b are both finite.

[1] Some authors use the term "neighborhood" to mean what is here called "open neighborhood."

Proposition 2.8. The closed sets of X have the properties that

(a) X and the empty set $\varnothing$ are closed,
(b) an arbitrary intersection of closed sets is closed,
(c) any finite union of closed sets is closed.

PROOF. This result follows from Proposition 2.5 by taking complements. In (a), the complements of X and $\varnothing$ are $\varnothing$ and X, respectively. For (b) and (c), we use the formulas $\big(\bigcap_\alpha F_\alpha\big)^c = \bigcup_\alpha F_\alpha^c$ and $\big(\bigcup_\alpha F_\alpha\big)^c = \bigcap_\alpha F_\alpha^c$ for the complements of intersections and unions.

If A is a subset of X, then x in X is a **limit point** of X if each neighborhood of x contains a point of A distinct from x. The **closure**[2] A^{cl} of A is the union of A with the set of all limit points of A. For example, the limit points of the set $[a, b) \cup \{b + 1\}$ on the real line are the points of the closed interval $[a, b]$, and the closure of the set is $[a, b] \cup \{b + 1\}$.

Proposition 2.9. A subset A of X is closed if and only if it contains all its limit points.

PROOF. Suppose A is closed, so that A^c is open. If x is in A^c, then A^c is an open neighborhood of x disjoint from A, so that x cannot be a limit point of A. Thus all limit points of A lie in A. In the reverse direction suppose that A contains all its limit points. If x is in A^c, then x is not a limit point of A, and hence there exists an open neighborhood of x lying completely in A^c. Since x is arbitrary, A^c is open, and thus A is closed.

Proposition 2.10. The closure A^{cl} of a subset A of X is closed. The closure of A is the intersection of all closed sets containing A; that is, it is the smallest closed set containing A.

PROOF. We shall apply Proposition 2.9. If x is given as a limit point of A^{cl}, we are to see that x is in A^{cl}. Assume the contrary. Then x is not in A, and x is not a limit point of A. Because of the latter condition, there exists an open neighborhood U of x that does not meet A except possibly in x. Because of the former condition, U does not meet A at all. Since x is a limit point of A^{cl}, U contains a point y of A^{cl}. Since U does not meet A, y has to be a limit point of A. Since U is an open neighborhood of y, U has to contain a point of A, and we have a contradiction. We conclude that x is in A^{cl}, and Proposition 2.9 shows that A^{cl} is closed.

Any closed set F containing A contains all its limit points, by Proposition 2.9, and hence contains all the limit points of A. Thus $F \supseteq A^{\mathrm{cl}}$. Since A^{cl} itself is a closed set containing A, it follows that A^{cl} is the smallest closed set containing A.

[2] Some authors write $\overline{A}$ instead of A^{cl} for the closure of A.

Corollary 2.11. A subset A of X is closed if and only if $A = A^{\mathrm{cl}}$. Consequently $(A^{\mathrm{cl}})^{\mathrm{cl}} = A^{\mathrm{cl}}$ for any subset A of X.

Two remarks are in order. The first remark is that the proofs of all the results from Proposition 2.6 through Corollary 2.11 use only that the family of open subsets of X satisfies properties (a), (b), and (c) in Proposition 2.5 and do not actually depend on the precise definition of "open set." This observation will be of importance to us in Chapter X, when properties (a), (b), and (c) will be taken as an axiomatic definition of a "topology" of open sets for X, and then all the results from Proposition 2.6 through Corollary 2.11 will still be valid.

The second remark is that the mathematics of pseudometric spaces can always be reduced to the mathematics of metric spaces, and we shall normally therefore work only with metric spaces. The device for this reduction is given in the next proposition, which uses the notion of an equivalence relation. Equivalence relations are taken as known but are reviewed in Section A6 of the appendix.

Proposition 2.12. Let (X, d) be a pseudometric space. If members x and y of X are called equivalent whenever $d(x, y) = 0$, then the result is an equivalence relation. Denote by $[x]$ the equivalence class of x and by X_0 the set of all equivalence classes. The definition $d_0([x], [y]) = d(x, y)$ consistently defines a function $d_0 : X_0 \times X_0 \to \mathbb{R}$, and (X_0, d_0) is a metric space. A subset A is open in X if and only if two conditions are satisfied: A is a union of equivalence classes, and the set A_0 of such classes is an open subset of X_0.

PROOF. The reflexive, symmetric, and transitive properties of the relation "equivalent" are immediate from the defining properties of a metric. Let x and x' be equivalent, and let y and y' be equivalent. Then

$$d(x, y) \le d(x, x') + d(x', y') + d(y', y) = 0 + d(x', y') + 0 = d(x', y'),$$

and similarly

$$d(x', y') \le d(x, y).$$

Thus $d(x, y) = d(x', y')$, and d_0 is well defined. The properties showing that d_0 is a metric are immediate from the corresponding properties for d.

Next let x be in an open set A, and let x' be equivalent to x. Since A is open, some open ball $B(r; x)$ is contained in A. Since x' has $d(x, x') = 0$, x' lies in $B(r; x)$. Thus x' lies in A, and A is the union of equivalence classes.

Finally let A be any union of equivalence classes, and let A_0 be the set of those classes. If x is in A, then the set of points in some equivalence class lying in $B(r; [x])$ is just $B(r; x)$, and it follows that A is open in X if and only if A_0 is open in X_0.

3. Continuous Functions

Before we discuss continuous functions between metric spaces, let us take note of some properties of inverse images for abstract functions as listed in Section A1 of the appendix. If $f : X \to Y$ is a function between two sets X and Y and E is a subset of Y, we denote by $f^{-1}(E)$ the inverse image of E under f, i.e., $\{x \in X \mid f(x) \in E\}$. The properties are that inverse images of functions respect unions, intersections, and complements.

Let (X, d) and (Y, ρ) be metric spaces. A function $f : X \to Y$ is **continuous at a point** $x \in X$ if for each $\epsilon > 0$, there is a $\delta > 0$ such that $\rho(f(x), f(y)) < \epsilon$ whenever $d(x, y) < \delta$. This definition is consistent with the definition when (X, d) and (Y, ρ) are both equal to $\mathbb{R}$ with the usual metric.

Proposition 2.13. If (X, d) and (Y, ρ) are metric spaces, then a function $f : X \to Y$ is continuous at the point $x \in X$ if and only if for any open neighborhood V of $f(x)$ in Y, there is a neighborhood U of x such that $f(U) \subseteq V$.

PROOF. Let f be continuous at x and let V be given. Choose $\epsilon > 0$ such that $B(\epsilon; f(x))$ is contained in V, and choose $\delta > 0$ such that $\rho(f(x), f(y)) < \epsilon$ whenever $d(x, y) < \delta$. Then $y \in B(\delta; x)$ implies $f(y) \in B(\epsilon; f(x)) \subseteq V$. Thus $U = B(\delta; x)$ has $f(U) \subseteq V$.

Conversely suppose that f satisfies the condition in the statement of the proposition. Let $\epsilon > 0$ be given, and choose a neighborhood U of x such that $f(U) \subseteq B(\epsilon; f(x))$. Since U is a neighborhood of x, we can find an open ball $B(\delta; x)$ lying in U. Then $f(B(\delta; x)) \subseteq B(\epsilon; f(x))$, and hence $\rho(f(x), f(y)) < \epsilon$ whenever $d(x, y) < \delta$.

Corollary 2.14. Let $f : X \to Y$ and $g : Y \to Z$ be functions between metric spaces. If f is continuous at x and g is continuous at $f(x)$, then the composition $g \circ f$, given by $(g \circ f)(y) = g(f(y))$, is continuous at x.

PROOF. Let W be an open neighborhood of $g(f(x))$. By continuity of g at $f(x)$, we can choose a neighborhood V of $f(x)$ such that $g(V) \subseteq W$. Possibly by passing to a subset of V, we may assume that V is an open neighborhood of $f(x)$. By continuity of f at x, we can choose a neighborhood U of x such that $f(U) \subseteq V$. Then $g(f(U)) \subseteq W$. Taking Proposition 2.13 into account, we see that $g \circ f$ is continuous at x.

Proposition 2.15. If (X, d) and (Y, ρ) are metric spaces and f is a function from X into Y, then the following are equivalent:

(a) the function f is continuous at every point of X,
(b) the inverse image under f of every open set in Y is open in X,
(c) the inverse image under f of every closed set in Y is closed in X.

PROOF. Suppose (a) holds. If V is open in Y and x is in $f^{-1}(V)$, then $f(x)$ is in V. Since f is continuous at x by (a), Proposition 2.13 gives us a neighborhood U of x, which we may take to be open, such that $f(U) \subseteq V$. Then we have $x \in U \subseteq f^{-1}(V)$. Since x is arbitrary in $f^{-1}(V)$, $f^{-1}(V)$ is open. Thus (b) holds. In the reverse direction, suppose (b) holds. Let x in X be given, and let V be an open neighborhood of $f(x)$. By (b), $U = f^{-1}(V)$ is open, and U is then an open neighborhood of x mapping into V. This proves (a), and thus (a) and (b) are equivalent. Conditions (b) and (c) are equivalent, since $f^{-1}(V)^c = f^{-1}(V^c)$.

A function $f : X \to Y$ that is continuous at every point of X, as in Proposition 2.15, will simply be said to be **continuous**. A function $f : X \to Y$ is a **homeomorphism** if f is continuous, if f is one-one and onto, and if $f^{-1} : Y \to X$ is continuous. The relation "is homeomorphic to" is an equivalence relation. Namely, the identity function shows that the relation is reflexive, the symmetry of the relation is built into the definition, and the transitivity follows from Corollary 2.14.

If (X, d) is a metric space and if A is a nonempty subset of X, then the **distance** from x to A, denoted by $D(x, A)$, is defined by

$$D(x, A) = \inf_{y \in A} d(x, y).$$

Proposition 2.16. Let A be a fixed nonempty subset of a metric space (X, d). Then the real-valued function f defined on X by $f(x) = D(x, A)$ is continuous.

PROOF. If x and y are in X and z is in A, then the triangle inequality gives

$$D(x, A) \leq d(x, z) \leq d(x, y) + d(y, z).$$

Taking the infimum over z gives $D(x, A) \leq d(x, y) + D(y, A)$. Reversing the roles of x and y, we obtain $D(y, A) \leq d(x, y) + D(x, A)$, since $d(y, x) = d(x, y)$. Therefore

$$|f(x) - f(y)| = |D(x, A) - D(y, A)| \leq d(x, y).$$

Fix x, let $\epsilon > 0$ be given, and take $\delta = \epsilon$. If $d(x, y) < \delta = \epsilon$, then our inequality gives us $|f(x) - f(y)| < \epsilon$. Hence f is continuous at x. Since x is arbitrary, f is continuous.

Corollary 2.17. If (X, d) is a metric space, then the real-valued function $d(x, y)$ for fixed x is continuous in y.

PROOF. This is the special case of the proposition in which A is the set $\{y\}$.

Corollary 2.18. Let (X, d) be a metric space, and let x be in X. Then the closed ball $\{y \in X \mid d(x, y) \leq r\}$ is a closed set.

REMARK. Nevertheless, the closed ball is not necessarily the closure of the open ball $B(r; x) = \{y \in X \mid d(x, y) < r\}$. A counterexample is provided by any open ball of radius 1 in a space with the discrete metric.

PROOF. If $f(y) = d(x, y)$, the set in question is $f^{-1}([0, r])$. Corollary 2.17 says that f is continuous, and the equivalence of (a) and (c) in Proposition 2.15 shows that the set in question is closed.

Proposition 2.19. If A is a nonempty subset of a metric space (X, d), then $A^{\mathrm{cl}} = \{x \mid D(x, A) = 0\}$.

PROOF. The set $\{x \mid D(x, A) = 0\}$ is closed by Propositions 2.16 and 2.15, and it contains A. By Proposition 2.10 it contains A^{cl}. For the reverse inclusion, suppose x is not in A^{cl}, hence that x is not in A and x is not a limit point of A. These conditions imply that there is some $\epsilon > 0$ such that $B(\epsilon; x)$ is disjoint from A, hence that $d(x, y) \geq \epsilon$ for all y in A. Taking the infimum over y gives $D(x, A) \geq \epsilon > 0$. Hence $D(x, A) \neq 0$.

4. Sequences and Convergence

For a set S, we have already defined in Section I.1 the notion of a **sequence** in S as a function from a certain kind of subset of integers into S. In this section we work with sequences in metric spaces.

A sequence $\{x_n\}$ in a metric space (X, d) is **eventually in** a subset A of X if there is an integer N such that x_n is in A whenever $n \geq N$. The sequence $\{x_n\}$ **converges** to a point x in X if the sequence is eventually in each neighborhood of x. It is apparent that if $\{x_n\}$ converges to x, then so does every subsequence $\{x_{n_k}\}$.

Proposition 2.20. If (X, d) is a metric space, then no sequence in X can converge to more than one point.

PROOF. Suppose on the contrary that $\{x_n\}$ converges to distinct points x and y. The number $m = d(x, y)$ is then > 0. By the assumed convergence, x_n lies in both open balls $B(\frac{m}{2}; x)$ and $B(\frac{m}{2}; y)$ if n is large enough. Thus x_n lies in the intersection of these balls. But this intersection is empty, since the presence of a point z in both balls would mean that $d(x, y) \leq d(x, z) + d(z, y) < \frac{m}{2} + \frac{m}{2} = m$, contradiction.

If a sequence $\{x_n\}$ in a metric space (X, d) converges to x, we shall call x the **limit** of the sequence and write $\lim_{n\to\infty} x_n = x$ or $\lim_n x_n = x$ or $\lim x_n = x$ or $x_n \to x$. A sequence has at most one limit, by Proposition 2.20. If the definition of convergence is extended to pseudometric spaces, then sequences need not have unique limits.

Let us identify convergent sequences in some of the examples of metric spaces in Section 1.

EXAMPLES OF CONVERGENCE IN METRIC SPACES.

(0) The real line. On $\mathbb{R}$ with the usual metric, the convergent sequences are the sequences convergent in the usual sense of Section I.1.

(1) Euclidean space $\mathbb{R}^n$. Here the metric is given by

$$d(x, y) = \Big(\sum_{k=1}^{n} (x_k - y_k)^2\Big)^{1/2}$$

if $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$. Another metric $d'(x, y)$ is given by

$$d'(x, y) = \max_{1\le k\le n} |x_k - y_k|,$$

and we readily check that

$$d'(x, y) \le d(x, y) \le \sqrt{n}\, d'(x, y).$$

From this inequality it follows that the convergent sequences in $(\mathbb{R}^n, d)$ are the same as the convergent sequences in $(\mathbb{R}^n, d')$. On the other hand, the definition of d' as a maximum means that we have convergence in $(\mathbb{R}^n, d')$ if and only if we have ordinary convergence in each entry. Thus convergence of a sequence of vectors in $(\mathbb{R}^n, d)$ means convergence in the k^{th} entry for all k with $1 \le k \le n$.

(2) Complex Euclidean space $\mathbb{C}^n$. As a metric space, $\mathbb{C}^n$ gets identified with $\mathbb{R}^{2n}$. Thus a sequence of vectors in $\mathbb{C}^n$ converges if and only if it converges entry by entry.

(3) Extended real line $\mathbb{R}^*$. Here the metric is given by $d(x, y) = |f(x) - f(y)|$ with $f(x) = x/(1 + |x|)$ if x is in $\mathbb{R}$, $f(-\infty) = -1$, and $f(+\infty) = +1$. We saw in Section 1 that the intersections with $\mathbb{R}$ of the open balls of $\mathbb{R}^*$ are the open intervals in $\mathbb{R}$. Thus convergence of a sequence in $\mathbb{R}^*$ to a point x in $\mathbb{R}$ means that the sequence is eventually in $(-\infty, +\infty)$ and thereafter is an ordinary convergent sequence in $\mathbb{R}$. Convergence to $+\infty$ of a sequence $\{x_n\}$ means that for each real number M, there is an integer N such that $x_n \ge M$ whenever $n \ge N$. Convergence to $-\infty$ is analogous.

(4) Bounded scalar-valued functions on S in the uniform metric. A sequence $\{f_n\}$ in $B(S)$ converges in the uniform metric on $B(S)$ if and only if $\{f_n\}$ converges uniformly, in the sense below, to some member f of $B(S)$. The definition of **uniform convergence** here is the natural generalization of the one in Section I.3: $\{f_n\}$ converges to f uniformly if for each $\epsilon > 0$, there is an integer N such that $n \geq N$ implies $|f_n(s) - f(s)| < \epsilon$ for all s simultaneously. An important fact in this case is that the sequence $\{f_n\}$ is **uniformly bounded**, i.e., that there exists a real number M such that $|f_n(s)| \leq M$ for all n and s. In fact, choose some integer N for $\epsilon = 1$. Then the triangle inequality gives

$$|f_n(s)| \leq |f_n(s) - f_N(s)| + |f_N(s)| \leq 1 + |f_N(s)|$$

for all s if $n \geq N$, so that M can be taken to be $\max_{1 \leq n \leq N} \{\sup_{s \in S} |f_n(s)|\} + 1$.

(5) Bounded functions from S into a metric space (R, ρ). Convergence here is the expected generalization of **uniform convergence**: $\{f_n\}$ converges to f uniformly if for each $\epsilon > 0$, there is an integer N such that $n \geq N$ implies $\rho(f_n(s), f(s)) < \epsilon$ for all s simultaneously. As in Example 4, a uniformly convergent sequence of bounded functions is **uniformly bounded** in the sense that $\rho(f_n(s), r_0) \leq M$ for all n and s, M being some real number. Here r_0 is any fixed member of R.

(7) Indiscrete space X. The function $d(x, y)$ in this case is a pseudometric, not a metric, unless X has only one point. Every sequence in X converges to every point in X.

(8) Discrete metric. Convergence of a sequence $\{x_n\}$ in a space X with the discrete metric means that $\{x_n\}$ is eventually constant.

(11) Hilbert cube. For each n, let $(\{x_m\}_{m=1}^{\infty})_n$ be a member of the Hilbert cube, and write x_{mn} for the m^{th} term of the n^{th} sequence. As n varies, the sequence of sequences converges if and only if $\lim_n x_{mn}$ exists for each m.

(12) L^1 metric on Riemann integrable functions. The function $d(f, g)$ defined in this case is a pseudometric, not a metric. Convergence in the corresponding metric space as in Proposition 2.12 therefore really means a certain kind of convergence of equivalence classes: If $\{f_n\}$ and f are given, the sequence of classes $\{[f_n]\}$ converges to the class $[f]$ if and only if $\lim_n \int_a^b |f_n(x) - f(x)|\,dx = 0$. The use of classes in the notation is rather cumbersome and not very helpful, and consequently it is common practice to treat the L^1 space as a metric space and to work with its members as if they were functions rather than equivalence classes. We return to this point in Chapter V.

Let us elaborate a little on Examples 4 and 5, concerning the space $B(S)$ of bounded scalar-valued functions on a set S or, more generally, the space of bounded functions from S into a metric space (R, ρ). Suppose that S has the additional structure of a metric space (S, d). We let $C(S)$ be the subset of $B(S)$ consisting of bounded continuous functions on S, and we write $C(S, \mathbb{R})$ or $C(S, \mathbb{C})$ if we want to be explicit about the range. More generally we consider the space of bounded continuous functions from S into the metric space R. All of these are metric spaces in their own right.

Proposition 2.21. Let (S, d) and (R, ρ) be metric spaces, let x_0 be in S, and let $f_n : S \to R$ be a sequence of bounded functions from S into R that converge uniformly to $f : S \to R$ and are continuous at x_0. Then f is continuous at x_0. In particular, the uniform limit of continuous functions is continuous.

PROOF. For x in S, we write

$$\rho(f(x), f(x_0)) \le \rho(f(x), f_n(x)) + \rho(f_n(x), f_n(x_0)) + \rho(f_n(x_0), f(x_0)).$$

Given $\epsilon > 0$, we choose an integer N by the uniform convergence such that the first and third terms on the right side are $< \epsilon$ for $n \ge N$. With N fixed, we choose $\delta > 0$ by the continuity of f_N at x_0 such that $\rho(f_N(x), f_N(x_0)) < \epsilon$ whenever $d(x, x_0) < \delta$. Then the displayed inequality shows that $d(x, x_0) < \delta$ implies $\rho(f(x), f(x_0)) < 3\epsilon$, and the proposition follows.

We conclude this section with some elementary results involving convergence of sequences in metric spaces.

Proposition 2.22. If (X, d) is a metric space, then

(a) for any subset A of X and limit point x of A, there exists a sequence in $A - \{x\}$ converging to x,

(b) any convergent sequence in X with limit $x \in X$ either has infinite image, with x as a limit point of the image, or else is eventually constantly equal to x.

REMARK. This result and the first corollary below are used frequently—and often without specific reference.

PROOF OF (a). For each $n \ge 1$, the open ball $B(1/n; x)$ is an open neighborhood of x and must contain a point x_n of A distinct from the limit point x. Then $d(x_n, x) < 1/n$, and thus $\lim x_n = x$. Hence $\{x_n\}$ is the required sequence.

PROOF OF (b). Suppose that $\{x_n\}$ converges to x and has infinite image. By discarding the terms equal to x, we obtain a subsequence $\{x_{n_k}\}$ with limit x. If U is an open neighborhood of x, then $\{x_{n_k}\}$ is eventually in U, by the assumed convergence. Since no term of the subsequence equals x, U contains a member of the image of $\{x_n\}$ different from x. Thus x is a limit point of the image of $\{x_n\}$.

Now suppose that $\{x_n\}$ converges to x and has finite image $\{p_1, \dots, p_r\}$. If x_n is equal to some particular p_{j_0} for infinitely many n, then $\{x_n\}$ has an infinite subsequence converging to p_{j_0}. Since $\{x_n\}$ converges to x, every convergent subsequence converges to x. Therefore $p_{j_0} = x$. For $j \neq j_0$, only finitely many x_n can then equal p_j, and it follows that $\{x_n\}$ is eventually constantly equal to $p_{j_0} = x$.

Corollary 2.23. If (X, d) is a metric space, then a subset F of X is closed if and only if every convergent sequence in F has its limit in F.

PROOF. Suppose that F is closed and $\{x_n\}$ is a convergent sequence in F with limit x. By Proposition 2.22b, either x is in the image of the sequence or x is a limit point of the sequence. In either case, x is in F; thus the limit of any convergent sequence in F is in F.

Conversely suppose every convergent sequence in F has its limit in F. If x is a limit point of F, then Proposition 2.22a produces a sequence in $F - \{x\}$ converging to x. By assumption, the limit x is in F. Therefore F contains all its limit points and is closed.

Corollary 2.24. If (S, d) is a metric space, then the set $C(S)$ of bounded continuous scalar-valued functions on S is a closed subset of the metric space $B(S)$ of all bounded scalar-valued functions on S.

PROOF. Proposition 2.21 shows for any sequence in $C(S)$ convergent in $B(S)$ that the limit is actually in $C(S)$. By Corollary 2.23, $C(S)$ is closed in $B(S)$.

Proposition 2.25. Let $f : X \to Y$ be a function between metric spaces. Then f is continuous at a point x in X if and only if whenever $\{x_n\}$ is a convergent sequence in X with limit x, then $\{f(x_n)\}$ is convergent in Y with limit $f(x)$.

REMARK. In the special case of domain and range $\mathbb{R}$, this result was mentioned in Section I.1 after the definition of continuity. We deferred the proof of the special case until now to avoid repetition.

PROOF. Suppose that f is continuous at x and that $\{x_n\}$ is a convergent sequence in X with limit x. Let V be any open neighborhood of $f(x)$. By continuity, there exists an open neighborhood U of x such that $f(U) \subseteq V$. Since $x_n \to x$, there exists N such that x_n is in U whenever $n \geq N$. Then $f(x_n)$ is in $f(U) \subseteq V$ whenever $n \geq N$. Hence $\{f(x_n)\}$ converges to $f(x)$.

Conversely suppose that $x_n \to x$ always implies $f(x_n) \to f(x)$. We are to show that f is continuous. Let V be an open neighborhood of $f(x)$. We are to show that some open neighborhood of x maps into V under f. Assuming the contrary, we can find, for each $n \geq 1$, some x_n in $B(1/n; x)$ such that $f(x_n)$ is not in V. Then $x_n \to x$, but the distance of $f(x_n)$ from $f(x)$ is bounded away from 0. Thus $f(x_n)$ cannot converge to $f(x)$. This is a contradiction, and we conclude that some $B(1/n; x)$ maps into V under f; since V is arbitrary, f is continuous.

5. Subspaces and Products

When working with functions on the real line, one frequently has to address situations in which the domain of the function is just an open interval or a closed interval, rather than the whole line. When one uses the ϵ-δ definition of continuity, the subject does not become much more cumbersome, but it can become more cumbersome if one uses some other definition, such as one involving limits. The theory of metric spaces has a device for addressing smaller domains than the whole space—the notion of a subspace—and then the theory of functions on a subspace stands on an equal footing with the theory of functions on the whole space.

Let (X, d) be a metric space, and let A be a nonempty subset of X. There is a natural way of making A into a metric space, namely by taking the restriction $d\big|_{A\times A}$ as a metric for A. When we do so, we speak of A as a **subspace** of X. When there is a need to be more specific, we may say that A is a **metric subspace** of X. If A is an open subset of X, we may say that A is an **open subspace**; if A is a closed subset of X, we may say that A is a **closed subspace**.

Proposition 2.26. If A is a subspace of a metric space (X, d), then the open sets of A are exactly all sets $U \cap A$, where U is open in X, and the closed sets of A are all sets $F \cap A$, where F is closed in X.

PROOF. The open balls in A are the intersections with A of the open balls of X, and the statement about open sets follows by taking unions. The closed sets of A are the complements within A of all the open sets of A, thus all sets of the form $A-(U\cap A)$ with U open in X. Since $A-(U\cap A) = A\cap U^c$, the statement about closed sets follows.

Corollary 2.27. If A is a subspace of (X, d) and if $f : X \to Y$ is continuous at a point a of A, then the restriction $f\big|_A$, mapping A into Y, is continuous at a. Also, f is continuous at a if and only if the function $f_0 : X \to f(X)$ obtained by redefining the range to be the image is continuous at a.

PROOF. Let V be an open neighborhood of $f(a)$ in Y. By continuity of f at a as a function on X, choose an open neighborhood U of a in X with $f(U) \subseteq V$. Then $U \cap A$ is an open neighborhood of a in A, and $f(U \cap A) \subseteq V$. Hence $f\big|_A$ is continuous at a.

The most general open neighborhood of $f(a)$ in $f(X)$ is of the form $V \cap f(X)$ with V an open neighborhood of $f(a)$ in Y. Since $f^{-1}(V) = f_0^{-1}(V \cap f(X))$, the condition for continuity of f_0 at a is the same as the condition for continuity of f at a.

We now turn our attention to product spaces. Product spaces are a convenient device for considering functions of several variables.

If (X, d) and (Y, d') are metric spaces, there are several natural ways of making the product set $X \times Y$, the set of ordered pairs with the first member from X and the second from Y, into a metric space, but all such ways lead to the same class of open sets and therefore also the same class of convergent sequences. We discussed an instance of this phenomenon in Example 1 of Section 4. For general X and Y, three such metrics on $X \times Y$ are

$$\begin{aligned}
\rho_1\big((x_1, y_1), (x_2, y_2)\big) &= d(x_1, x_2) + d'(y_1, y_2),\\
\rho_2\big((x_1, y_1), (x_2, y_2)\big) &= \big(d(x_1, x_2)^2 + d'(y_1, y_2)^2\big)^{1/2},\\
\rho_\infty\big((x_1, y_1), (x_2, y_2)\big) &= \max\{d(x_1, x_2),\ d'(y_1, y_2)\}.
\end{aligned}$$

Each satisfies the defining properties of a metric. Simple algebra gives

$$\max\{a, b\} \le (a^2 + b^2)^{1/2} \le a + b \le 2\max\{a, b\}$$

whenever a and b are nonnegative reals, and therefore

$$\rho_\infty \le \rho_2 \le \rho_1 \le 2\rho_\infty.$$

Let us check that this chain of inequalities implies that the neighborhoods of a point (x_0, y_0) are the same in all three metrics, hence that the open sets are the same in all three metrics. For any $r > 0$, the open balls about (x_0, y_0) in the three metrics satisfy

$$B_1(r; (x_0, y_0)) \subseteq B_2(r; (x_0, y_0)) \subseteq B_\infty(r; (x_0, y_0)) \subseteq B_1(2r; (x_0, y_0)).$$

The first and second inclusions show that open balls about (x_0, y_0) in the metrics ρ_2 and ρ_∞ are neighborhoods of (x_0, y_0) in the metric ρ_1. Similarly the second and third inclusions show that open balls in the metrics ρ_∞ and ρ_1 are neighborhoods in the metric ρ_2, and the third and first inclusions show that open balls in the metrics ρ_1 and ρ_2 are neighborhoods in the metric ρ_∞.

We shall refer to the metric ρ_∞ as the **product metric** for $X \times Y$. If $X \times Y$ is being regarded as a metric space and no metric has been mentioned, ρ_∞ is to be understood. But it is worth keeping in mind that ρ_1 and ρ_2 yield the same open sets. In the case of Euclidean space, it is the metric ρ_2 on $\mathbb{R}^m \times \mathbb{R}^n$ that gives the

Euclidean metric on $\mathbb{R}^{m+n}$; thus the product metric and the Euclidean metric are distinct but yield the same open sets.

A sequence $\{(x_n, y_n)\}$ in the product metric converges to (x_0, y_0) in $X \times Y$ if and only if $\{x_n\}$ converges to x_0 and $\{y_n\}$ converges to y_0. Since the three metrics on $X \times Y$ yield the same convergent sequences, this statement is valid in the metrics ρ_1 and ρ_2 as well.

It is an elementary property of the arithmetic operations in $\mathbb{R}$ that if $\{x_n\}$ converges to x_0 and $\{y_n\}$ converges to y_0, then $\{x_n + y_n\}$ converges to $x_0 + y_0$. Similar statements apply to subtraction, multiplication, maximum, and minimum, and then to absolute value and to division except where division by 0 is involved. Further similar statements apply to those operations on vectors that make sense. Applying Proposition 2.25, we obtain (a) through (e) in the following proposition. Conclusions (a$'$) through (e$'$) are proved similarly.

Proposition 2.28. The following operations are continuous:

(a) addition and subtraction from $\mathbb{R}^n \times \mathbb{R}^n$ into $\mathbb{R}^n$,
(b) scalar multiplication from $\mathbb{R} \times \mathbb{R}^n$ into $\mathbb{R}^n$,
(c) the map $x \mapsto x^{-1}$ from $\mathbb{R} - \{0\}$ to $\mathbb{R} - \{0\}$,
(d) the map $x \mapsto |x|$ from $\mathbb{R}^n$ to $\mathbb{R}$,
(e) the operations from $\mathbb{R}^2$ to $\mathbb{R}$ of taking the maximum of two real numbers and taking the minimum of two real numbers,
(a$'$) addition and subtraction from $\mathbb{C}^n \times \mathbb{C}^n$ into $\mathbb{C}^n$,
(b$'$) scalar multiplication from $\mathbb{C} \times \mathbb{C}^n$ into $\mathbb{C}^n$,
(c$'$) the map $x \mapsto x^{-1}$ from $\mathbb{C} - \{0\}$ to $\mathbb{C} - \{0\}$,
(d$'$) the map $x \mapsto |x|$ from $\mathbb{C}^n$ to $\mathbb{R}$,
(e$'$) the map $x \mapsto \bar{x}$ from $\mathbb{C}$ to $\mathbb{C}$.

Corollary 2.29. Let (X, d) be a metric space, and let f and g be continuous functions from X into $\mathbb{R}^n$ or $\mathbb{C}^n$. If c is a scalar, then $f+g$, cf, $f-g$, and $|f|$ are continuous. If $n = 1$, then the product fg is continuous, and the function $1/f$ is continuous on the set where f is not zero. If $n = 1$ and the functions take values in $\mathbb{R}$, then $\max\{f, g\}$ and $\min\{f, g\}$ are continuous. If $n = 1$ and the functions take values in $\mathbb{C}$, then the complex conjugate $\bar{f}$ is continuous.

REMARKS. If (S, d) is a metric space, then it follows that the metric space $C(S)$ of bounded continuous scalar-valued functions on S is a vector space. As such, it is a vector subspace of the metric space $B(S)$ of bounded scalar-valued functions on S, and it is a metric subspace as well.[3]

[3]The word "subspace" can now be used in two senses, that of a metric subspace of a metric space and that of a vector subspace of a vector space. The latter kind of subspace we shall always refer to as a "vector subspace," retaining the word "vector" for clarity. A "closed vector subspace" of $B(S)$ then has to mean a closed metric subspace that is also a vector subspace.

PROOF. The argument for $f+g$ and for functions with values in $\mathbb{R}^n$ will illustrate matters sufficiently. We set up $x \mapsto f(x)+g(x)$ as a suitable composition, expressing the composition in a diagram:

$$X \xrightarrow{x \mapsto (x,x)} X \times X \xrightarrow{(x,y)\mapsto(f(x),g(y))} \mathbb{R}^n \times \mathbb{R}^n \xrightarrow{(u,v)\mapsto u+v} \mathbb{R}^n.$$

Each function in the diagram is continuous, the last of them by Proposition 2.28a, and then the composition is continuous by Corollary 2.14.

We conclude this section with one further remark. When (X, d) is a metric space, we saw in Corollary 2.17 that $x \mapsto d(x, y)$ and $y \mapsto d(x, y)$ are continuous functions from X to $\mathbb{R}$. Actually, $(x, y) \mapsto d(x, y)$ is a continuous function from $X \times X$ into $\mathbb{R}$ if we use the product metric. In fact, if ρ_∞ denotes the product metric with $\rho_\infty\big((x, y), (x_0, y_0)\big) = \max\{d(x, x_0), d(y, y_0)\}$, then we have $d(x, y) \le d(x, x_0) + d(x_0, y_0) + d(y_0, y)$ and therefore

$$d(x, y) - d(x_0, y_0) \le d(x, x_0) + d(y, y_0).$$

Reversing the roles of (x, y) and (x_0, y_0), we see that

$$\begin{aligned} |d(x, y) - d(x_0, y_0)| &\le d(x, x_0) + d(y, y_0) \\ &\le 2\max\{d(x, x_0), d(y, y_0)\} \\ &= 2\rho_\infty\big((x, y), (x_0, y_0)\big). \end{aligned}$$

From this chain of inequalities, it follows that d is continuous with $\delta = \epsilon/2$.

6. Properties of Metric Spaces

This section contains two results about metric spaces. One lists a number of "separation properties" of sets within any metric space. The other concerns the completely different property of "separability," which is satisfied by some metric spaces and not by others, and it says that separability may be defined in any of three equivalent ways.

Proposition 2.30 (separation properties). Let (X, d) be a metric space. Then

(a) every one-point subset of X is a closed set, i.e., X is $\mathbf{T}_1$,

(b) for any two distinct points x and y of X, there are disjoint open sets U and V with $x \in U$ and $y \in V$, i.e., X is **Hausdorff**,

(c) for any point $x \in X$ and any closed set $F \subseteq X$ with $x \notin F$, there are disjoint open sets U and V with $x \in U$ and $F \subseteq V$, i.e., X is **regular**,

(d) for any two disjoint closed subsets E and F of X, there are disjoint open sets U and V such that $E \subseteq U$ and $F \subseteq V$, i.e., X is **normal**,

(e) for any two disjoint closed subsets E and F of X, there is a continuous function $f : X \to [0, 1]$ such that f is 0 exactly on E and f is 1 exactly on F.

PROOF. For (a), the set $\{x\}$ is the intersection of all closed balls $B(r; x)$ for $r > 0$ and hence is closed by Corollary 2.18 and Proposition 2.8b. For (e), the function $f(x) = D(x; E)/(D(x; E) + D(x; F))$ is continuous by Proposition 2.16 and Corollary 2.29 and takes on the values 0 and 1 exactly on E and F, respectively, by Proposition 2.19.

For (d), we need only apply (e) and Proposition 2.15b with $U = f^{-1}\big((-\infty, \frac{1}{2})\big)$ and $V = f^{-1}\big((\frac{1}{2}, +\infty)\big)$. Conclusions (a) and (d) imply (c), and conclusions (a) and (c) imply (b). This completes the proof.

A **base** $\mathcal{B}$ for a metric space (X, d) is a family of open sets such that every open set is a union of members of $\mathcal{B}$. The family of all open balls is an example of a base.

Proposition 2.31. If (X, d) is a metric space, then a family $\mathcal{B}$ of subsets of X is a base for (X, d) if and only if

(a) every member of $\mathcal{B}$ is open and
(b) for each $x \in X$ and open neighborhood U of x, there is some member B of $\mathcal{B}$ such that x is in B and B is contained in U.

PROOF. If $\mathcal{B}$ is a base, then (a) holds by definition of base. If U is open in X, then $U = \bigcup_\alpha B_\alpha$ for some members B_α of $\mathcal{B}$, and any such B_α containing x can be taken as the set B in (b).

Conversely suppose that $\mathcal{B}$ satisfies (a) and (b). By (a), each member of $\mathcal{B}$ is open in X. If U is open in X, we are to show that U is a union of members of $\mathcal{B}$. For each $x \in U$, choose some set $B = B_x$ as in (b). Then $U = \bigcup_{x \in U} B_x$, and hence each open set in X is a union of members of $\mathcal{B}$. Thus $\mathcal{B}$ is a base.

This book uses the word **countable** to mean finite or countably infinite. It is then meaningful to ask whether a particular metric space (X, d) has a countable base. On the real line $\mathbb{R}$, the open intervals with rational endpoints form a countable base.

A subset D of X is **dense in** a subset A of X if $D^{\mathrm{cl}} \supseteq A$; D is **dense**, or **everywhere dense**, if D is dense in X. A set D is dense if and only if there is some point of D in each nonempty open set of X.

A family $\mathcal{U}$ of open sets is an **open cover** of X if the union of the sets in $\mathcal{U}$ is X. An **open subcover** of $\mathcal{U}$ is a subfamily of $\mathcal{U}$ that is itself an open cover.

Proposition 2.32. The following three conditions are equivalent for a metric space (X, d):

(a) X has a countable base,
(b) every open cover of X has a countable open subcover,
(c) X has a countable dense subset.

PROOF. If (a) holds, let $\mathcal{B} = \{B_n\}_{n\geq 1}$ be a countable base, and let $\mathcal{U}$ be an over cover of X. Any $U \in \mathcal{U}$ is the union of the $B_n \in \mathcal{B}$ with $B_n \subseteq U$. If $\mathcal{B}_0 = \{B_n \in \mathcal{B} \mid B_n \subseteq U \text{ for some } U \in \mathcal{U}\}$, then it follows that $\bigcup_{B_n \in \mathcal{B}_0} = \bigcup_{U \in \mathcal{U}} = X$. For each B_n in $\mathcal{B}_0$, select some U_n in $\mathcal{U}$ with $B_n \subseteq U_n$. Then $\bigcup_n U_n \supseteq \bigcup_{B_n \in \mathcal{B}_0} = X$, and $\{U_n\}$ is a countable open subcover of $\mathcal{U}$. Thus (b) holds.

If (b) holds, form, for each fixed $n \geq 1$, the open cover of X consisting of all open balls $B(1/n; x)$. For that n, let $\{B(1/n; x_{mn})\}_{m\geq 1}$ be a countable open subcover. We shall prove that the set D of all x_{mn}, with m and n arbitrary, is dense in X. It is enough to prove that each nonempty open set in X contains a member of D, hence to prove, for each n, that each open ball of radius $1/n$ contains a member of D. Thus consider $B(1/n; x)$. Since the open balls $B(1/n; x_{mn})$ with $m \geq 1$ cover X, x is in some $B(1/n; x_{mn})$. Then that x_{mn} has $d(x_{mn}, x) < 1/n$, and hence x_{mn} is in $B(1/n; x)$. Thus D is dense, and (c) holds.

If (c) holds, let $\{x_n\}_{n\geq 1}$ be a countable dense set. Form the collection of all open balls centered at some x_n and having rational radius. Let us use Proposition 2.31 to see that this collection of open sets, which is certainly countable, is a base. Let U be an open neighborhood of x. We are to see that there is some member B of our collection such that x is in B and B is contained in U. Since U is a neighborhood of x, we can find an open ball $B(r; x)$ such that $B(r; x) \subseteq U$; we may assume that r is rational. The given set $\{x_n\}_{n\geq 1}$ being dense, some x_n lies in $B(r/2; x)$. If y is in $B(r/2; x_n)$, then $d(x, y) \leq d(x, x_n) + d(x_n, y) < \frac{r}{2} + \frac{r}{2} = r$. Hence x lies in $B(r/2; x_n)$ and $B(r/2; x_n) \subseteq B(r; x) \subseteq U$. Since $r/2$ is rational, the open ball $B(r/2; x_n)$ is in our countable collection, and our countable collection is a base. This proves (a).

A metric space satisfying the equivalent conditions of Proposition 2.32 is said to be **separable**. Among the examples of metric spaces in Section 1, the ones in Examples 1, 2, 3, 6, 8 if X is countable, 9, 11, 12, 13, and 14 are separable. A countable dense set in Examples 1, 2, and 3 is given by all points with all coordinates rational. In Example 6, one countable dense set consists of all sequences with only finitely many nonzero entries, those being rational, and in Examples 8 and 9, X itself is a countable dense set. In Example 11, the sequences that are 0 in all but finitely many entries, those being rational, form a countable dense set. In Example 13, the set of finite linear combinations of exponentials e^{inx} using scalars in $\mathbb{Q} + i\mathbb{Q}$ is dense as a consequence of Parseval's equality. In Example 12, when $[a, b] = [-\pi, \pi]$, the same countable set as for Example 13 is dense by Proposition 2.25 because the sets of functions in Examples 12 and 13 coincide and the inclusion of L^2 into L^1 is continuous. In Example 14, the set of polynomials with coefficients in $\mathbb{Q}+i\mathbb{Q}$ is countable and can be shown to be dense.

Example 10 is not separable, and Example 8 is not separable if X is uncountable.

7. Compactness and Completeness

In Section 6 we introduced the notions of open cover and subcover for a metric space. We call a metric space **compact** if every open cover of the space has a *finite* subcover. A subset E of a metric space (X, d) is **compact** if it is compact as a subspace of the whole space, i.e., if every collection of open sets in X whose union contains E has a finite subcollection whose union contains E.

Historically this notion was embodied in the Heine–Borel Theorem, which says that any closed bounded subset of Euclidean space has the property that has just been defined to be compactness. As we shall see in Theorem 2.36 and Corollary 2.37 below, the Heine–Borel Theorem can be proved from the Bolzano–Weierstrass Theorem (Theorem 1.8) and leads to faster, more transparent proofs of some of the consequences of the Bolzano–Weierstrass Theorem. Even more important is that it generalizes beyond metric spaces and produces useful conclusions about certain spaces of functions when statements about pointwise convergence of a sequence of functions are inadequate.

Easily established examples of compact sets are hard to come by. For one example, consider in a metric space (X, d) a convergent sequence $\{x_n\}$ along with its limit x. The subset $E = \{x\} \cup \bigcup_n \{x_n\}$ of X is compact. In fact, if $\mathcal{U}$ is an open cover of E, some member U of $\mathcal{U}$ has x as an element, and then all but finitely many elements of the sequence must be in U as well. Say that U contains x and all x_n with $n \geq N$. For $1 \leq n < N$, let U_n be a member of $\mathcal{U}$ containing x_n. Then $\{U, U_1, \ldots, U_{N-1}\}$ is a finite subcover of $\mathcal{U}$.

It is easier to exhibit noncompact sets. The open interval $(0, 1)$ is not compact, as is seen from the open cover $\{(\frac{1}{n}, 1)\}$. Nor is an infinite discrete space, since one-point sets form an open cover. A subtle dramatic example is the closed unit ball C of the hedgehog space X, Example 10 in Section 1; this set is not compact. In fact, the open ball of radius $1/2$ about the origin is an open set in X, and so is each open ray from the origin out to infinity. Let $\mathcal{U}$ be this collection of open sets. Then $\mathcal{U}$ is an open cover of C. However, no member of $\mathcal{U}$ is superfluous, since for each U in $\mathcal{U}$, there is some point x in C such that x is in C but x is in no other member of $\mathcal{U}$. Thus $\mathcal{U}$ does not contain even a countable subcover.

Let us now work directly toward a proof of the equivalence of compactness and the Bolzano–Weierstrass property in a metric space.

Proposition 2.33. A compact metric space is separable.

PROOF. This is immediate from equivalent condition (b) for the definition of separability in Proposition 2.32.

Proposition 2.34. In any metric space (X, d),

(a) every compact subset is closed and bounded and
(b) any closed subset of a compact set is compact.

PROOF. For (a), let E be a compact subset of X, fix x_0 in X, and let U_n for $n \geq 1$ be the open ball $\{x \in X \mid d(x_0, x) < n\}$. Then $\{U_n\}$ is an open cover of E. Since the U_n's are nested, the compactness of E implies that E is contained in a single U_N for some N. Then every member of E is at distance at most N from x_0, and E is bounded.

To see that E is closed, we argue by contradiction. Let x_0' be a limit point of E that is not in E. By the Hausdorff property (Proposition 2.30b), we can find, for each $x \in E$, open sets U_x and V_x with $x \in U_x$, $x_0' \in V_x$, and $U_x \cap V_x = \varnothing$. The sets U_x form an open cover of E. By compactness let $\{U_{x_1}, \ldots, U_{x_n}\}$ be a finite subcover. Then $E \subseteq U_{x_1} \cup \cdots \cup U_{x_n}$, which is disjoint from the neighborhood $V_{x_1} \cap \cdots \cap V_{x_n}$ of x_0'. Thus x_0' cannot be a limit point of E, and we have arrived at a contradiction. This proves (a).

For (b), let E be compact, and let F be a closed subset of E. Because of (a), F is a closed subset of X. Let $\mathcal{U}$ be an open cover of F. Then $\mathcal{U} \cup \{F^c\}$ is an open cover of E. Passing to a finite subcover and discarding F^c, we obtain a finite subcover of F. Thus F is compact.

A collection of subsets of a nonempty set is said to have the **finite-intersection property** if each intersection of finitely many of the subsets is nonempty.

Proposition 2.35. A metric space (X, d) is compact if and only if each collection of closed subsets of X with the finite-intersection property has nonempty intersection.

PROOF. Closed sets with the finite-intersection property have complements that are open sets, no finite subcollection of which is an open cover.

Theorem 2.36. A metric space (X, d) is compact if and only if every sequence has a convergent subsequence.

PROOF. Suppose that X is compact. Arguing by contradiction, suppose that $\{x_n\}_{n\geq 1}$ is a sequence in X with no convergent subsequence. Put $F = \bigcup_{n=1}^{\infty}\{x_n\}$. The subset F of X is closed by Corollary 2.23, hence compact by Proposition 2.34b. Since no x_n is a limit point of F, there exists an open set U_n in X containing x_n but no other member of F. Then $\{U_n\}_{n\geq 1}$ is an open cover of F with no finite subcover, and we have arrived at a contradiction.

Conversely suppose that every sequence has a convergent subsequence. We first show that X is separable. Fix an integer n. There cannot be infinitely many disjoint open balls of radius $1/n$, since otherwise we could find a sequence from among their centers with no convergent subsequence. Thus we can choose a finite disjoint collection of these open balls that is not contained in a larger such finite collection. Let their centers be $x_1, \ldots, x_N$. The claim is that every point of X is at distance $< 2/n$ from one of these finitely many centers. In fact, if $x \in X$ is given, form $B(\frac{1}{n}; x)$. This must meet some $B(\frac{1}{n}; x_i)$ at a point y, and then

$$d(x, x_i) \leq d(x, y) + d(y, x_i) < \tfrac{1}{n} + \tfrac{1}{n} = \tfrac{2}{n}.$$

Thus x is at distance $< 2/n$ from one of the finitely many centers, as asserted. Now let n vary, and let D be the set of all these centers for all n. Then every point of X has members of D arbitrarily close to it, and hence D is a countable dense set in X. Thus X is separable.

Let $\mathcal{U}$ be an open cover of X having no finite subcover. By the separability and condition (b) in Proposition 2.32, we may assume that $\mathcal{U}$ is countable, say $\mathcal{U} = \{U_1, U_2, \dots\}$. Since $U_1 \cup U_2 \cup \cdots \cup U_n$ is not a cover, there exists a point x_n not in the union of the first n sets. By hypothesis the sequence $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$, say with limit x. Since $\mathcal{U}$ is a cover, some member U_N of $\mathcal{U}$ contains x. Then $\{x_{n_k}\}$ is eventually in U_N, and some n_k with $n_k > N$ has x_{n_k} in U_N. But x_{n_k} is not in $U_1 \cup \cdots \cup U_{n_k}$ by construction, and this union contains U_N, since $n_k > N$. We have arrived at a contradiction, and we conclude that $\mathcal{U}$ must have had a finite subcover.

Corollary 2.37 (Heine–Borel Theorem) In Euclidean space $\mathbb{R}^n$, every closed bounded set is compact.

REMARK. Conversely we saw in Proposition 2.34a that every compact subset of any metric space is closed and bounded.

PROOF. Let C be a closed rectangular solid in $\mathbb{R}^n$, and let $x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})$ be the members of a sequence in C. By the Bolzano–Weierstrass Theorem (Theorem 1.8) for $\mathbb{R}^1$, we can find a subsequence convergent in the first coordinate, a subsequence of that convergent in the second coordinate, and so on. Thus $\{x^{(k)}\}$ has a convergent subsequence. By Theorem 2.36, C is compact. Applying Corollary 2.34b, we see that every closed bounded subset of $\mathbb{R}^n$ is compact.

The next few results will show how the use of compactness both simplifies and generalizes some of the theorems proved in Section I.1.

Proposition 2.38. Let (X, d) and (Y, ρ) be metric spaces with X compact. If $f : X \to Y$ is continuous, then $f(X)$ is a compact subset of Y.

PROOF. If $\{U_\alpha\}$ is an open cover of $f(X)$, then $\{f^{-1}(U_\alpha)\}$ is an open cover of X. Let $\{f^{-1}(U_j)\}_{j=1}^n$ be a finite subcover. Then $\{U_j\}_{j=1}^n$ is a finite subcover of $f(X)$.

Corollary 2.39. Let (X, d) be a compact metric space, and let $f : X \to \mathbb{R}$ be a continuous function. Then f attains its maximum and minimum values.

REMARK. Theorem 1.11 was the special case of this result with $X = [a, b]$. This particular space X is compact by the Heine–Borel Theorem (Corollary 2.37), and the corollary applies to yield exactly the conclusion of Theorem 1.11.

PROOF. By Proposition 2.38, $f(X)$ is a compact subset of $\mathbb{R}$. By Proposition 2.34a, $f(X)$ is closed and bounded. The supremum and infimum of the members of $f(X)$ in $\mathbb{R}^*$ lie in $\mathbb{R}$, since $f(X)$ is bounded, and they are limits of sequences in $f(X)$. Since $f(X)$ is closed, Proposition 2.23 shows that they must lie in $f(X)$.

Corollary 2.40. Let (X, d) and (Y, ρ) be metric spaces with X compact. If $f : X \to Y$ is continuous, one-one, and onto, then f is a homeomorphism.

REMARK. In the hypotheses of the change of variables formula for integrals in $\mathbb{R}^1$ (Theorem 1.34), a function $\varphi : [A, B] \to [a, b]$ was given as strictly increasing, continuous, and onto. Another hypothesis of the theorem was that φ^{-1} was continuous. Corollary 2.40 shows that this last hypothesis was redundant.

PROOF. Let E be a closed subset of X, and consider $(f^{-1})^{-1}(E) = f(E)$. The set E is compact by Proposition 2.34b, $f(E)$ is compact by Proposition 2.38, and $f(E)$ is closed by Proposition 2.34a. Proposition 2.15b thus shows that f^{-1} is continuous.

If (X, d) and (Y, ρ) are metric spaces, a function $f : X \to Y$ is **uniformly continuous** if for each $\epsilon > 0$, there is some $\delta > 0$ such that $d(x_1, x_2) < \delta$ implies $\rho(f(x_1), f(x_2)) < \epsilon$. This is the natural generalization of the definition in Section I.1 for the special case of a real-valued function of a real variable.

Proposition 2.41. Let (X, d) and (Y, ρ) be metric spaces with X compact. If $f : X \to Y$ is continuous, then f is uniformly continuous.

REMARK. This result generalizes Theorem 1.10, which is the special case $X = [a, b]$ and $Y = \mathbb{R}$.

PROOF. Let $\epsilon > 0$ be given. For each $x \in X$, choose $\delta_x > 0$ such that $d(x', x) < \delta_x$ implies $\rho(f(x'), f(x)) < \epsilon/2$. The open balls $B(\frac{1}{2}\delta_x; x)$ cover X; let the balls with centers $x_1, \dots, x_n$ be a finite subcover. Put $\delta = \frac{1}{2}\min\{\delta_{x_1}, \dots, \delta_{x_n}\}$. Now suppose that $d(x', x) < \delta$. The point x is in some ball in the finite subcover; suppose x is in $B(\frac{1}{2}\delta_{x_j}; x_j)$. Then $d(x, x_j) < \frac{1}{2}\delta_{x_j}$, so that

$$d(x', x_j) \le d(x', x) + d(x, x_j) < \delta + \tfrac{1}{2}\delta_{x_j} \le \delta_{x_j}.$$

By definition of δ_{x_j}, $\rho(f(x'), f(x_j)) < \epsilon/2$ and $\rho(f(x_j), f(x)) < \epsilon/2$. Therefore

$$\rho(f(x'), f(x)) \le \rho(f(x'), f(x_j)) + \rho(f(x_j), f(x)) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,$$

and the proof is complete.

One final application of compactness is the Fundamental Theorem of Algebra, which is discussed in Section A8 of the appendix in the context of properties of polynomials.

Theorem 2.42 (Fundamental Theorem of Algebra). Every polynomial with complex coefficients and degree ≥ 1 has a complex root.

PROOF. Let $P : \mathbb{C} \to \mathbb{C}$ be the function $P(z) = \sum_{j=0}^{n} a_j z^j$, where $a_0, \dots, a_n$ are in $\mathbb{C}$ with $a_n \neq 0$ and with $n \geq 1$. We may assume that $a_n = 1$. Let $m = \inf_{z \in \mathbb{C}} |P(z)|$. Since $P(z) = z^n\big(1 + a_{n-1}z^{-1} + \cdots + a_1 z^{-(n-1)} + a_0 z^{-n}\big)$, we have $\lim_{z\to\infty} P(z)/z^n = 1$. Thus there exists an R such that $|P(z)| \geq \frac{1}{2}|z|^n$ whenever $|z| \geq R$. Choosing $R = R_0$ such that $\frac{1}{2}R_0^n \geq 2m$, we see that $|P(z)| \geq 2m$ for $|z| \geq R_0$. Consequently $m = \inf_{|z|\leq R_0} |P(z)|$. The set $S = \big\{z \in \mathbb{C} \,\big|\, |z| \leq R_0\big\}$ is compact by the Heine–Borel Theorem (Corollary 2.37), and Corollary 2.39 shows that $|P(z)|$ attains its minimum on S at some point z_0 in S. Then $|P(z)|$ attains its minimum on $\mathbb{C}$ at z_0. We shall show that this minimum value m is 0.

Assuming the contrary, define $Q(z) = P(z + z_0)/P(z_0)$, so that $Q(z)$ is a polynomial of degree $n \geq 1$ with $Q(0) = 1$ and $|Q(z)| \geq 1$ for all z. Write

$$Q(z) = 1 + b_k z^k + b_{k+1} z^{k+1} + \cdots + b_n z^n \qquad \text{with } b_k \neq 0.$$

Corollary 1.45 produces a real number θ such that $e^{ik\theta} b_k = -|b_k|$. For any $r > 0$ with $r^k |b_k| < 1$, we then have

$$\big|1 + b_k r^k e^{ik\theta}\big| = 1 - r^k |b_k|.$$

For such r and that θ, this equality implies that

$$\begin{aligned} |Q(z)| &\leq \big|1 + b_k r^k e^{ik\theta}\big| + r^{k+1}|b_{k+1}| + \cdots + r^n |b_n| \\ &\leq 1 - r^k\big(|b_k| - r|b_{k+1}| - \cdots - r^{n-k}|b_n|\big). \end{aligned}$$

For sufficiently small $r > 0$, the expression in parentheses on the right side is positive, and then $|Q(re^{i\theta})| < 1$, in contradiction to hypothesis. Thus we must have had $m = 0$, and we obtain $P(z_0) = 0$.

Another theme discussed in Section I.1 is that Cauchy sequences in $\mathbb{R}^1$ are convergent. This convergence was proved in Theorem 1.9 as a consequence of the Bolzano–Weierstrass Theorem. Actually, many sequences in metric spaces of importance in analysis are shown to converge without one's knowing the limit in advance and without using any compactness, and we therefore isolate the forced convergence of Cauchy sequences as a definition. In a metric space (X, d), a sequence $\{x_n\}$ is a **Cauchy sequence** if for any $\epsilon > 0$, there is some integer N

such that $d(x_m, x_n) < \epsilon$ whenever m and n are $\geq N$. A familiar 2ϵ argument shows that convergent sequences are Cauchy. Other familiar arguments show that any Cauchy sequence with a convergent subsequence is convergent and that any Cauchy sequence is bounded.

We say that the metric space (X, d) is **complete** if every Cauchy sequence in X converges to a point in X. We know that the line $\mathbb{R}^1$ is complete. It follows that $\mathbb{R}^n$ is complete because a Cauchy sequence in $\mathbb{R}^n$ is Cauchy in each coordinate. A nonempty subset E of X is **complete** if E as a subspace is a complete metric space. The next two propositions and corollary give three examples of complete metric spaces.

Proposition 2.43. A subset E of a complete metric space X is complete if and only if it is closed.

REMARK. In particular every closed subset of $\mathbb{R}^n$ is a complete metric space.

PROOF. Suppose E is closed. Let $\{x_n\}$ be a Cauchy sequence in E. Then $\{x_n\}$ is Cauchy in X, and the completeness of X implies that $\{x_n\}$ converges, say to some $x \in X$. By Corollary 2.23, x is in E. Thus $\{x_n\}$ is convergent in E. The converse is immediate from Corollary 2.23.

Proposition 2.44. If S is a nonempty set, then the vector space $B(S)$ of bounded scalar-valued functions on S, with the uniform metric, is a complete metric space.

PROOF. Let $\{f_n\}$ be a Cauchy sequence in $B(S)$. Then $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{C}$ for each x in S. Define $f(x) = \lim_n f_n(x)$. For any $\epsilon > 0$, we know that there is an integer N such that $|f_n(x) - f_m(x)| < \epsilon$ whenever n and m are $\geq N$. Taking into account the continuity of the distance function on $\mathbb{C}$, i.e., the continuity of absolute value, we let m tend to infinity and obtain $|f_n(x) - f(x)| \leq \epsilon$ for $n \geq N$. Thus $\{f_n\}$ converges to f in $B(S)$.

Corollary 2.45. Let (S, d) be a metric space. Then the vector space $C(S)$ of bounded continuous scalar-valued functions on S, with the uniform metric, is a complete metric space.

REMARK. $C(S)$ was observed to be a vector subspace in the remarks with Corollary 2.29.

PROOF. The space $B(S)$ is complete by Proposition 2.44, and $C(S)$ is a closed metric subspace by Corollary 2.24. Then $C(S)$ is complete by Proposition 2.43.

Now we shall relate compactness and completeness. A metric space (X, d) is said to be **totally bounded** if for any $\epsilon > 0$, finitely many open balls of radius ϵ cover X.

Theorem 2.46. A metric space (X, d) is compact if and only if it is totally bounded and complete.

PROOF. Let (X, d) be compact. If $\epsilon > 0$ is given, the open balls $B(\epsilon; x)$ cover X. By compactness some finite number of the balls cover X. Therefore X is totally bounded. Next let a Cauchy sequence $\{x_n\}$ be given. By Theorem 2.36, $\{x_n\}$ has a convergent subsequence. A Cauchy sequence with a convergent subsequence is necessarily convergent, and it follows that X is complete.

In the reverse direction, let X be totally bounded and complete. Theorem 2.36 shows that it is enough to prove that any sequence $\{x_n\}$ in X has a convergent subsequence. By total boundedness, find finitely many open balls of radius 1 covering X. Then infinitely many of the x_n's have to lie in one of these balls, and hence there is a subsequence $\{x_{n_k}\}$ that lies in a single one of these balls of radius 1. Next finitely many open balls of radius 1/2 cover X. In the same way there is a subsequence $\{x_{n_{k_l}}\}$ of $\{x_{n_k}\}$ that lies in a single one of these balls of radius 1/2. Continuing in this way, we can find successive subsequences, the m^{th} of which lies in a single ball of radius $1/m$. The Cantor diagonal process, used in the proof of Theorem 1.22, allows us to form a single subsequence $\{x_{j_i}\}$ of $\{x_n\}$ such that for each m, $\{x_{j_i}\}$ is eventually in a ball of radius 2^{-m}. If $\epsilon > 0$ is given, find m such that $2^{-m} < \epsilon$, and let c_m be the center of the ball of radius $1/m$. Choose an integer N such that x_{j_i} lies in $B(1/m; c_m)$ whenever $j_i \geq N$. If $j_i \geq N$ and $j_{i'} \geq N$, then $d(c_m, x_{j_i}) < \epsilon$ and $d(c_m, x_{j_{i'}}) < \epsilon$, whence $d(x_{j_i}, x_{j_{i'}}) < 2\epsilon$. Therefore the subsequence $\{x_{j_i}\}$ is Cauchy. By completeness it converges. Hence $\{x_n\}$ has a convergent subsequence, and the theorem is proved.

Let (X, d) and (Y, ρ) be metric spaces, and let $f : X \to Y$ be uniformly continuous. Then f carries Cauchy sequences to Cauchy sequences. In fact, if $\{x_n\}$ is Cauchy in X and if $\epsilon > 0$ is given, choose some δ of uniform continuity for f and ϵ, and find an integer N such that $d(x_n, x_{n'}) < \delta$ whenever n and n' are $\geq N$. Then $\rho(f(x_n), f(x_{n'})) < \epsilon$ for the same n's and n''s, and hence $\{f(x_n)\}$ is Cauchy.

Proposition 2.47. Let (X, d) and (Y, ρ) be metric spaces with Y complete, let D be a dense subset of X, and let $f : D \to Y$ be uniformly continuous. Then f extends uniquely to a continuous function $F : X \to Y$, and F is uniformly continuous.

PROOF OF UNIQUENESS. If x is in X, apply Proposition 2.22a to choose a sequence $\{x_n\}$ in D with $x_n \to x$. Continuity of F forces $F(x_n) \to F(x)$. But $F(x_n) = f(x_n)$ for all n. Thus $F(x) = \lim_n f(x_n)$ is forced.

PROOF OF EXISTENCE. If x is in X, choose $x_n \in D$ with $x_n \to x$. Since $\{x_n\}$ is convergent, it is Cauchy. Since f is uniformly continuous, $\{f(x_n)\}$ is

Cauchy. The completeness of Y then allows us to define $F(x) = \lim f(x_n)$, but we must see that F is well defined. For this purpose, suppose also that $\{y_n\}$ is a sequence in D that converges to x. Let $\{z_n\}$ be the sequence $x_1, y_1, x_2, y_2, \ldots$. This sequence is Cauchy, and $\{x_n\}$ and $\{y_n\}$ are subsequences of it. Therefore $\lim f(y_n) = \lim f(z_n) = \lim f(x_n)$, and $F(x)$ is well defined.

For the uniform continuity of F, let $\epsilon > 0$ be given, and choose some δ of uniform continuity for f and $\epsilon/3$. Suppose that x and x' are in X with $d(x, x') < \delta/3$. Choose x_n in D with $d(x_n, x) < \delta/3$ and $\rho(f(x_n), F(x)) < \epsilon/3$, and choose x'_n in D with $d(x'_n, x') < \delta/3$ and $\rho(f(x'_n), F(x')) < \epsilon/3$. Then $d(x_n, x'_n) < \delta$ by the triangle inequality, and hence $\rho(f(x_n), f(x'_n)) < \epsilon/3$. Thus $\rho(F(x), F(x')) < \epsilon$ by the triangle inequality.

8. Connectedness

Although the Intermediate Value Theorem (Theorem 1.12) in Section I.1 was derived from the Bolzano–Weierstrass Theorem, the Intermediate Value Theorem is not to be regarded as a consequence of compactness. Instead, the relevant property is "connectedness," which we discuss in this section.

A metric space (X, d) is **connected** if X cannot be written as $X = U \cup V$ with U and V open, disjoint, and nonempty. A subset E of X is **connected** if E is connected as a subspace of X, i.e., if E cannot be written as a disjoint union $(E \cap U) \cup (E \cap V)$ with U and V open in X and with $E \cap U$ and $E \cap V$ both nonempty. The disjointness in this definition is of $E \cap U$ and $E \cap V$; the open sets U and V may have nonempty intersection.

Proposition 2.48. The connected subsets of $\mathbb{R}$ are the intervals—open, closed, and half open.

PROOF. Let E be a connected subset of $\mathbb{R}$, and suppose that there are real numbers a, b, c such that $a < c < b$, a and b are in E, and c is not in E. Forming the open sets $U = (-\infty, c)$ and $V = (c, +\infty)$ in $\mathbb{R}$, we see that E is the disjoint union of $E \cap U$ and $E \cap V$ and that these two sets are nonempty. Thus E is not connected.

Conversely suppose that I is an open, closed, or half-open interval of $\mathbb{R}$ from a to b, with $a \neq b$ but with a or b or both allowed to be infinite. Arguing by contradiction, suppose that I is not connected. Choose open sets U and V in $\mathbb{R}$ such that I is the disjoint union of $I \cap U$ and $I \cap V$ and these two sets are nonempty. Without loss of generality, there exist members c and c' of $I \cap U$ and $I \cap V$, respectively, with $c < c'$. Since U is open and c has to be $< b$, all real numbers $c + \epsilon$ with $\epsilon > 0$ sufficiently small are in $I \cap U$. Let $d = \sup \big\{x \,\big|\, [c, u) \subseteq I \cap U\big\}$, so that $d > c$.

If $d < b$, then the fact that U is open implies that d is not in $I \cap U$. Thus d is in $I \cap V$. Since V is open and $d > a$, $d - \epsilon$ is in $I \cap V$ if $\epsilon > 0$ is sufficiently small. But then $d - \epsilon$ is in both $I \cap U$ and $I \cap V$ for ϵ sufficiently small. This is a contradiction, and we conclude that $d = b$.

If $d = b$ is in $I \cap V$, then the same argument shows that $b - \epsilon$ is in both $I \cap U$ and $I \cap V$ for ϵ positive and sufficiently small, and we again have a contradiction. Consequently all points from c to the right end of I are in $I \cap U$. This is again a contradiction, since c' is known to be in $I \cap V$.

Proposition 2.49. The continuous image of a connected metric space is connected.

PROOF. Let (X, d) and (Y, ρ) be metric spaces, and let $f : X \to Y$ be continuous. We are to prove that $f(X)$ is connected. Corollary 2.27 shows that there is no loss of generality in assuming that $f(X) = Y$, i.e., f is onto. Arguing by contradiction, suppose that Y is the union $Y = U \cup V$ of disjoint nonempty open sets. Then $X = f^{-1}(U) \cup f^{-1}(V)$ exhibits X as the disjoint union of nonempty sets, and these sets are open as a consequence of Proposition 2.15a. Thus X is not connected.

Corollary 2.50 (Intermediate Value Theorem). For real-valued functions of a real variable, the continuous image of any interval is an interval.

PROOF. This is immediate from Propositions 2.48 and 2.49.

Further connected sets beyond those in $\mathbb{R}$ are typically built from other connected sets. One tool is a **path** in X, which is a continuous function from a closed bounded interval $[a, b]$ into X. The image of a path is connected by Propositions 2.48 and 2.49. A metric space (X, d) is **pathwise connected** if for any two points x_1 and x_2 in X, there is some path p from x_1 to x_2, i.e., if there is some continuous $p : [a, b] \to X$ with $p(a) = x_1$ and $p(b) = x_2$.

A pathwise-connected metric space (X, d) is necessarily connected. In fact, otherwise we could write X as a disjoint union of two nonempty open sets U and V. Let x_1 be in U and x_2 be in V, and let $p : [a, b] \to X$ be a path from x_1 to x_2. Then $p([a, b]) = (p([a, b]) \cap U) \cup (p([a, b]) \cap V)$ exhibits $p([a, b])$ as a disjoint union of relatively open sets, and these sets are nonempty, since x_1 is in the first set and x_2 is in the second set. Consequently $p([a, b])$ is not connected, in contradiction to the fact that the image of any path is connected.

We can view a pathwise-connected metric space as the union of images of paths from a single point to all other points, and such a union is then connected. The following proposition generalizes this construction.

Proposition 2.51. If (X, d) is a metric space and $\{E_\alpha\}$ is a system of connected subsets of X with a point x_0 in common, then $\bigcup_\alpha E_\alpha$ is connected.

PROOF. Assuming the contrary, find open sets U and V in X such that $\bigcup_\alpha E_\alpha$ is the disjoint union of its intersections with U and V and these two intersections are both nonempty. Say that x_0 is in U. Since E_α is connected and x_0 is in $E_\alpha \cap U$, the decomposition $E_\alpha = (E_\alpha \cap U) \cup (E_\alpha \cap V)$ forces $E_\alpha \cap V$ to be empty. Then $\big(\bigcup_\alpha E_\alpha\big) \cap V = \bigcup_\alpha (E_\alpha \cap V)$ is empty, and we have arrived at a contradiction.

Proposition 2.52. If (X, d) is a metric space and E is a connected subset of X, then the closure E^{cl} is connected.

PROOF. Suppose that U and V are open sets in X such that E^{cl} is contained in $U \cup V$ and $E^{\text{cl}} \cap U \cap V$ is empty. We are to prove that $E^{\text{cl}} \cap U$ and $E^{\text{cl}} \cap V$ cannot both be nonempty. Arguing by contradiction, let x be in $E^{\text{cl}} \cap U$ and let y be in $E^{\text{cl}} \cap V$. Since E is connected, $E \cap U$ and $E \cap V$ cannot both be nonempty, and thus x and y cannot both be in E. Thus at least one of them, say x, is a limit point of E. Since U is a neighborhood of x, U contains a point e of E different from x. Thus e is in $E \cap U$. Since y cannot then be in $E \cap V$, y is a limit point of E. Since V is a neighborhood of y, V contains a point f of E different from y. Thus f is in $E \cap V$, and we have arrived at a contradiction.

EXAMPLE. The graph in $\mathbb{R}^2$ of $\sin(1/x)$ for $0 < x \le 1$ is pathwise connected, and we have seen that pathwise-connected sets are connected. The closure of this graph consists of the graph together with all points $(0, t)$ for $-1 \le t \le 1$, and this closure is connected by Proposition 2.52. One can show, however, that this closure is not pathwise connected. Thus we obtain an example of a connected set in $\mathbb{R}^2$ that is not pathwise connected.

9. Baire Category Theorem

A number of deep results in analysis depend critically on the fact that some metric space is complete. Already we have seen that the metric space $C(S)$ of bounded continuous scalar-valued functions on a metric space is complete, and we shall see as not too hard a consequence in Chapter XII that there exists a continuous periodic function whose Fourier series diverges at a point. One of the features of the Lebesgue integral in Chapter V will be that the metric spaces of integrable functions and of square-integrable functions, with their natural metrics, are further examples of complete metric spaces. Thus these spaces too are available for applications that make use of completeness.

The main device through which completeness is transformed into a powerful hypothesis is the Baire Category Theorem below. A closed set in a metric space is **nowhere dense** if its interior is empty. Its complement is an open dense set, and conversely the complement of any open dense set is closed nowhere dense.

EXAMPLE. A nontrivial example of a closed nowhere dense set is a **Cantor set**[4] in $\mathbb{R}$. This is a set constructed from a closed bounded interval of $\mathbb{R}$ by removing an open interval in the middle of length a fraction r_1 of the total length with $0 < r_1 < 1$, removing from each of the 2 remaining closed subintervals an open interval in the middle of length a fraction r_2 of the total length of the subinterval, removing from each of the 4 remaining closed subintervals an open interval in the middle of length a fraction r_3 of the total length of the interval, and so on indefinitely. The Cantor set is obtained as the intersection of the approximating sets. It is closed, being the intersection of closed sets, and it is nowhere dense because it contains no interval of more than one point. For the **standard Cantor set**, the starting interval is $[0, 1]$, and the fractions are given by $r_1 = r_2 = \cdots = \frac{1}{3}$ at every stage. In general, the "length" of the resulting set[5] is the product of the length of the starting interval and $\prod_{n=1}^{\infty} (1 - r_n)$.

Theorem 2.53 (Baire Category Theorem). If (X, d) is a complete metric space, then

(a) the intersection of countably many open dense sets is nonempty,
(b) X is not the union of countably many closed nowhere dense sets.

PROOF. Conclusions (a) and (b) are equivalent by taking complements. Let us prove (a). Suppose that U_n is open and dense for $n \geq 1$. Since U_1 is nonempty and open, let E_1 be an open ball $B(r_1; x_1)$ whose closure is in U_1 and whose radius is $r_1 \leq 1$. We construct inductively open balls $E_n = B(r_n; x_n)$ with $r_n \leq \frac{1}{n}$ such that $E_n \subseteq U_1 \cap \cdots \cap U_n$ and $E_n^{\text{cl}} \subseteq E_{n-1}$. Suppose E_n with $n \geq 1$ has been constructed. Since U_{n+1} is dense and E_n is nonempty and open, $U_{n+1} \cap E_n$ is not empty. Let x_{n+1} be a point in $U_{n+1} \cap E_n$. Since $U_{n+1} \cap E_n$ is open, we can find an open ball $E_{n+1} = B(r_{n+1}; x_{n+1})$ with radius $r_{n+1} \leq \frac{1}{n+1}$ and center the point x_{n+1} in U_{n+1} such that $E_{n+1}^{\text{cl}} \subseteq U_{n+1} \cap E_n$. Then E_{n+1} has the required properties, and the inductive construction is complete. The sequence $\{x_n\}$ is Cauchy because whenever $n \geq m$, the points x_n and x_m are both in E_m and thus have $d(x_n, x_m) < \frac{1}{m}$. Since X is by assumption complete, let $x_n \to x$. For any integer N, the inequality $n > N$ implies that x_n is in E_{N+1}. Thus the limit x is in $E_{N+1}^{\text{cl}} \subseteq E_N \subseteq U_1 \cap \cdots \cap U_N$. Since N is arbitrary, x is in $\bigcap_{n=1}^{\infty} U_n$.

[4]Often a mathematician who refers to "the" Cantor set is referring to what is called the "standard Cantor set" later in the present paragraph.

[5]To be precise, the length is the "Lebesgue measure" of the set in the sense to be defined in Chapter V.

REMARK. In (a), the intersection in question is dense, not merely nonempty. To see this, we observe in the first part of the proof that since U_1 is dense, E_1 can be chosen to be arbitrarily close to any member of X and to have arbitrarily small radius. Following through the construction, we see that x is in E_1 and hence can be arranged to be as close as we want to any member of X. The corresponding conclusion in (b) is that a nonempty open subset of X is never contained in the countable union of closed nowhere dense sets.

EXAMPLES.

(1) The subset $\mathbb{Q}$ of rationals in $\mathbb{R}$ is not the countable intersection of open sets. In fact, assume the contrary, and write $\mathbb{Q} = \bigcap_{n=1}^{\infty} U_n$ with U_n open. Each set U_n contains $\mathbb{Q}$ and hence is dense in $\mathbb{R}$. Also, for $q \in \mathbb{Q}$, the set $\mathbb{R} - \{q\}$ is open and dense. Thus the equality $\mathbb{Q} = \bigcap_{n=1}^{\infty} U_n$ implies that

$$\Big(\bigcap_{n=1}^{\infty} U_n\Big) \cap \Big(\bigcap_{q \in \mathbb{Q}} (\mathbb{R} - \{q\})\Big)$$

is empty, in contradiction to Theorem 2.53.

(2) Let us start with a Cantor set as at the beginning of this section. The total interval is to be $[0, 1]$, and the set is to be built with middle segments of fractions $r_1, r_2, \ldots$. Within the closure of each removed open interval, we insert a Cantor set for that interval, possibly with different fractions $r_1, r_2, \ldots$ for each inserted Cantor set. This insertion involves further removed open intervals, and we insert a Cantor set into each of these. We continue this process indefinitely. The union of the constructed sets is dense. Can it be the entire interval $[0, 1]$? The answer is "no" because each of the Cantor sets is closed nowhere dense and because by Theorem 2.53, the interval $[0, 1]$ is not the countable union of closed nowhere dense sets.

A subset E of a metric space is said to be of the **first category** if it is contained in the countable union of closed nowhere dense sets. Theorem 2.53 and the remark after it together imply that no nonempty open set in a complete metric space is of the first category.

Theorem 2.54. Let (X, d) be a complete metric space, and let U be an open subset of X. Suppose for $n \geq 1$ that $f_n : U \to \mathbb{C}$ is a continuous function and that f_n converges pointwise to a function $f : U \to \mathbb{C}$. Then the set of discontinuities of f is of the first category.

The proof will make use of the notion of the **oscillation** of a function. For any function $g : U \to \mathbb{C}$, define

$$\operatorname{osc}_g(x_0) = \lim_{\delta \downarrow 0} \sup_{x \in B(\delta; x_0)} |g(x) - g(x_0)|,$$

so that g is continuous at x_0 if and only if $\operatorname{osc}_g(x_0) = 0$. At first glance it might seem that the sets $\{x \mid \operatorname{osc}_g(x) \geq r\}$ are always closed, no matter what discontinuities g has. Actually, these sets need not be closed. Take, for example, the function $g : \mathbb{R} \to \mathbb{R}$ that is 1 at every nonzero rational, 0 at every irrational, and 1/2 at 0. Then $\operatorname{osc}_g(x)$ is 1 at every x in $\mathbb{R}$ except for $x = 0$, where it is 1/2. Thus, in this example, the set $\{x \mid \operatorname{osc}_g(x) \geq 1\}$ is $\mathbb{R} - \{0\}$ and is not closed.

Lemma 2.55. Let (X, d) be a complete metric space, and let U be an open subset of X. If $g : U \to \mathbb{C}$ is a function and $\epsilon > 0$ is a positive number, then

$$\{x \in U \mid \operatorname{osc}_g(x) \geq 2\epsilon\}^{\text{cl}} \subseteq \{x \in U \mid \operatorname{osc}_g(x) \geq \tfrac{\epsilon}{2}\}.$$

PROOF. We need to see that the limit points of the set on the left are in the set on the right. Thus suppose that $\operatorname{osc}_g(x_n) \geq 2\epsilon$ for all n and that $x_n \to x_0$. For each n, choose $x_{n,m}$ such that $\lim_m x_{n,m} = x_n$ and $|g(x_{n,m}) - g(x_n)| \geq \epsilon$ for all m. Because of the convergence of $x_{n,m}$ to x_n, we may choose, for each n, an integer $m = m_n$ such that $d(x_{n,m_n}, x_n) < d(x_0, x_n)$, and then $\lim_n x_{n,m_n} = x_0$ by the triangle inequality. From $|g(x_{n,m_n}) - g(x_n)| \geq \epsilon$, the triangle inequality forces

$$|g(x_{n,m_n}) - g(x_0)| \geq \tfrac{\epsilon}{2} \qquad \text{or} \qquad |g(x_n) - g(x_0)| \geq \tfrac{\epsilon}{2}. \tag{$*$}$$

Defining y_n to be x_{n,m_n} or x_n according as the first or second inequality is the case in $(*)$, we have $y_n \to x_0$ and $|g(y_n) - g(x_0)| \geq \frac{\epsilon}{2}$. This proves the lemma.

PROOF OF THEOREM 2.54. In view of Lemma 2.55 and the fact that U is not of first category (Theorem 2.53 and the remark afterward), it is enough to prove for each $\epsilon > 0$ that $\{x \mid \operatorname{osc}_f(x) \geq \epsilon\}$ does not contain a nonempty open subset of X. Assuming the contrary, suppose that it contains the nonempty open set V. Define

$$A_{mn} = \{x \in V \mid |f_m(x) - f_n(x)| \leq \tfrac{\epsilon}{4}\}.$$

This is a relatively closed subset of V. Then $A_m = \bigcap_{n \geq m} A_{mn}$ is closed in V. If x is in V, the fact that $\{f_n(x)\}$ is a Cauchy sequence implies that there is some m such that x is in A_{mn} for all $n \geq m$. Hence $\bigcup_{m=1}^{\infty} A_m = V$. Again by Theorem 2.53 and the remark after it, some A_m has nonempty interior. Fix that m, and let W be its nonempty interior. Since

$$A_m \subseteq \{x \in V \mid |f_m(x) - f(x)| \leq \tfrac{\epsilon}{4}\},$$

every point of W has $|f_m(x) - f(x)| \leq \frac{\epsilon}{4}$ and $\operatorname{osc}_f(x) \geq \epsilon$. Let x_0 be in W and choose x_n tending to x_0 with $|f(x_n) - f(x_0)| \geq \frac{3\epsilon}{4}$. From $|f_m(x_n) - f(x_n)| \leq \frac{\epsilon}{4}$ and $|f_m(x_0) - f(x_0)| \leq \frac{\epsilon}{4}$, we obtain $|f_m(x_n) - f_m(x_0)| \geq \frac{\epsilon}{4}$. Since x_n converges to x_0, this inequality contradicts the continuity of f_m at x_0.

10. Properties of $C(S)$ for Compact Metric S

If (S, d) is a metric space, then we saw in Proposition 2.44 that the vector space $B(S)$ of bounded scalar-valued functions on S, in the uniform metric, is a complete metric space. We saw also in Corollary 2.45 that the vector subspace $C(S)$ of bounded continuous functions is a complete subspace. In this section we shall study the space $C(S)$ further under the assumption that S is compact. In this case Propositions 2.38 and 2.34 tell us that every continuous scalar-valued function on S is automatically bounded and hence is in $C(S)$.

The first result about $C(S)$ for S compact is a generalization of Ascoli's Theorem from its setting in Theorem 1.22 for real-valued functions on a bounded interval $[a, b]$. The generalized theorem provides an insight that is not so obvious from the special case that S is a closed bounded interval of $\mathbb{R}$. The insight is a characterization of the compact subsets of $C(S)$ when S is compact, and it is stated precisely in Corollary 2.57 below. The relevant definitions for Ascoli's Theorem are generalized in the expected way. Let $\mathcal{F} = \{f_\alpha \mid \alpha \in A\}$ be a set of scalar-valued functions on the compact metric space S. We say that $\mathcal{F}$ is **equicontinuous** at $x \in S$ if for each $\epsilon > 0$, there is some $\delta > 0$ such that $d(t, x) < \delta$ implies $|f(t) - f(x)| < \epsilon$ for all $f \in \mathcal{F}$. The set $\mathcal{F}$ of functions is **pointwise bounded** if for each $t \in [a, b]$, there exists a number M_t such that $|f(t)| \le M_t$ for all $f \in \mathcal{F}$. The set is **uniformly equicontinuous** on S if it is equicontinuous at each point $x \in S$ and if the δ can be taken independent of x. The set is **uniformly bounded** on S if it is pointwise bounded at each $t \in S$ and the bound M_t can be taken independent of t; this last definition is consistent with the definition of a uniformly bounded sequence of functions given in Section 4.

Theorem 2.56 (Ascoli's Theorem). Let (S, d) be a compact metric space. If $\{f_n\}$ is a sequence of scalar-valued functions on S that is equicontinuous at each point of S and pointwise bounded on S, then

(a) $\{f_n\}$ is uniformly equicontinuous and uniformly bounded on S,
(b) $\{f_n\}$ has a uniformly convergent subsequence.

REMARKS. The proof involves only notational changes from the special case Theorem 1.22; there are enough such changes, however, so that it is worth writing out the details. Inspection of this proof shows also that the range $\mathbb{R}$ or $\mathbb{C}$ may be replaced by any compact metric space. We shall see a further generalization of this theorem in Chapter X, and the proof at that time will look quite different.

PROOF. Since each f_n is continuous at each point, we know from Propositions 2.38, 2.34a, and 2.41 that each f_n is uniformly continuous and bounded. The proof of (a) amounts to an argument that the estimates in those theorems can be arranged to apply simultaneously for all n.

First consider the question of uniform boundedness. Choose, by Corollary 2.39, some x_n in S with $|f_n(x_n)|$ equal to $K_n = \sup_{x\in S}|f_n(x)|$. Then choose a subsequence on which the numbers K_n tend to $\sup_n K_n$ in $\mathbb{R}^*$. There will be no loss of generality in assuming that this subsequence is our whole sequence. By compactness of S, apply the Bolzano–Weierstrass property given in Theorem 2.36 to find a convergent subsequence $\{x_{n_k}\}$ of $\{x_n\}$, and let x_0 be the limit of this subsequence. By pointwise boundedness, find M_{x_0} with $|f_n(x_0)| \le M_{x_0}$ for all n. Then choose some δ of equicontinuity at x_0 for $\epsilon = 1$. As soon as k is large enough so that $d(x_{n_k}, x_0) < \delta$, we have

$$K_{n_k} = |f_{n_k}(x_{n_k})| \le |f_{n_k}(x_{n_k}) - f_{n_k}(x_0)| + |f_{n_k}(x_0)| < 1 + M_{x_0}.$$

Thus $1 + M_{x_0}$ is a uniform bound for the functions f_n.

For the uniform equicontinuity, fix $\epsilon > 0$. The uniform continuity of f_n for each n, as given in Proposition 2.41, means that it makes sense to define

$$\delta_n(\epsilon) = \min\left\{1,\ \sup\left\{\delta' > 0 \;\middle|\; \begin{array}{l} |f(x) - f(y)| < \epsilon \text{ whenever} \\ d(x,y) < \delta' \text{ and } x \text{ and } y \text{ are in } S \end{array}\right\}\right\}.$$

If $d(x,y) < \delta_n(\epsilon)$, then $|f_n(x) - f_n(y)| < \epsilon$. Put $\delta(\epsilon) = \inf_n \delta_n(\epsilon)$. Let us see that it is enough to prove that $\delta(\epsilon) > 0$: If x and y are in S with $d(x,y) < \delta(\epsilon)$, then $d(x,y) < \delta(\epsilon) \le \delta_n(\epsilon)$. Hence $|f_n(x) - f_n(y)| < \epsilon$ as required.

Thus we are to prove that $\delta(\epsilon) > 0$. If $\delta(\epsilon) = 0$, then we first choose a strictly increasing sequence $\{n_k\}$ of positive integers such that $\delta_{n_k}(\epsilon) < \frac{1}{k}$, and we next choose x_k and y_k in S with $d(x_k, y_k) < \delta_{n_k}(\epsilon)$ and $|f_k(x_k) - f_k(y_k)| \ge \epsilon$. Using the Bolzano–Weierstrass property again, we obtain a subsequence $\{x_{k_l}\}$ of $\{x_k\}$ such that $\{x_{k_l}\}$ converges, say to a limit x_0. Then

$$\limsup_l d(y_{k_l}, x_0) \le \limsup_l d(y_{k_l}, x_{k_l}) + \limsup_l d(x_{k_l}, x_0) = 0 + 0 = 0,$$

so that $\{y_{k_l}\}$ converges to x_0. Now choose, by equicontinuity at x_0, a number $\delta' > 0$ such that $|f_n(x) - f_n(x_0)| < \frac{\epsilon}{2}$ for all n whenever $d(x, x_0) < \delta'$. The convergence of $\{x_{k_l}\}$ and $\{y_{k_l}\}$ to x_0 implies that for large enough l, we have $d(x_{k_l}, x_0) < \delta'$ and $d(y_{k_l}, x_0) < \delta'$. Therefore $|f_{k_l}(x_{k_l}) - f_{k_l}(x_0)| < \frac{\epsilon}{2}$ and $|f_{k_l}(y_{k_l}) - f_{k_l}(x_0)| < \frac{\epsilon}{2}$, from which we conclude that $|f_{k_l}(x_{k_l}) - f_{k_l}(y_{k_l})| < \epsilon$. But we saw that $|f_k(x_k) - f_k(y_k)| \ge \epsilon$ for all k, and thus we have arrived at a contradiction. This proves the uniform equicontinuity and completes the proof of (a).

To prove (b), let R be a compact set containing all sets image(f_n). Choose a countable dense set D in S by Proposition 2.33. Using the Cantor diagonal process and the Bolzano–Weierstrass property of R, we construct a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ that is convergent at every point in D. Let us prove that $\{f_{n_k}\}$ is uniformly Cauchy. Let $\epsilon > 0$ be given, and let δ be some corresponding number exhibiting equicontinuity. The balls $B(\delta; r)$ centered at the members r of D cover S, and the compactness of S gives us finitely many of their centers $r_1, \dots, r_l$ such that any member of S is within δ of at least one of $r_1, \dots, r_l$. Then choose N with $|f_n(r_j) - f_m(r_j)| < \epsilon$ for $1 \le j \le l$ whenever n and m are $\ge N$. If x is in S, let $r(x)$ be an r_j with $d(x, r(x)) < \delta$. Whenever n and m are $\ge N$, we then have

$$\begin{aligned} &|f_n(x) - f_m(x)| \\ &\qquad \le |f_n(x) - f_n(r(x))| + |f_n(r(x)) - f_m(r(x))| + |f_m(r(x)) - f_m(x)| \\ &\qquad < \epsilon + \epsilon + \epsilon = 3\epsilon. \end{aligned}$$

Hence $\{f_{n_k}\}$ is uniformly Cauchy, and (b) follows since the metric space $C(S)$ is complete.

Corollary 2.57. If (S, d) is a compact metric space, then a subset E of $C(S)$ in the uniform metric has compact closure if and only if E is uniformly bounded and uniformly equicontinuous.

PROOF. First let us see that if E is uniformly bounded and uniformly equicontinuous, then so is E^{cl}. In fact, if $|f(x)| \le M$ for $f \in E$, then the same thing is true of any uniform limit of such functions. Hence E^{cl} is uniformly bounded. For the uniform equicontinuity of E^{cl}, let ϵ be given, and find some δ of equicontinuity for ϵ and the members of E. If f is a limit point of E, we can find a sequence $\{f_n\}$ in E converging uniformly to f. If $d(x, y) < \delta$, then the inequality

$$|f(x) - f(y)| \le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|$$

and the uniform convergence show that we obtain $|f(x) - f(y)| < 3\epsilon$ by fixing any sufficiently large n. Thus E^{cl} is uniformly equicontinuous.

Now suppose that E is a closed subset of $C(S)$ that is uniformly bounded and equicontinuous. Then Theorem 2.56 shows that any sequence in E has a subsequence that is convergent in $C(S)$. Since E is closed, the sequence is convergent in E. Theorem 2.36 then shows that E is compact.

Conversely suppose that E is compact in $C(S)$. Distance from 0 in $C(S)$ is a continuous real-valued function by Corollary 2.17, and this continuous function has to be bounded on the compact set E. Thus E is uniformly bounded. For the uniform equicontinuity, let $\epsilon > 0$ be given. Theorem 2.46 shows that E is totally bounded. Hence we can find a finite set $f_1, \dots, f_l$ in E such that each member f of E has $\sup_{x \in S} |f(x) - f_j(x)| < \epsilon$ for some j. By uniform continuity of each f_i, choose some number $\delta > 0$ such that $d(x, y) < \delta$ implies $|f_i(x) - f_i(y)| < \epsilon$ for $1 \le i \le l$. If f_j is the member of the finite set associated with f, then $d(x, y) < \delta$ implies

$$|f(x) - f(y)| \le |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_j(y) - f(y)| < 3\epsilon.$$

Hence E is uniformly equicontinuous.

The second result about $C(S)$ when S is compact generalizes the Weierstrass Approximation Theorem (Theorem 1.52) of Section I.9. We shall make use of a special case of the Weierstrass theorem in the proof—that $|x|$ is the uniform limit on $[-1, 1]$ of polynomials $P_n(x)$ with $P_n(0) = 0$. This special case was proved also by a direct argument in Section I.8.

Let us distinguish the case of real-valued functions from that of complex-valued functions, writing $C(S, \mathbb{R})$ and $C(S, \mathbb{C})$ in the two cases. The theorem in question gives a sufficient condition for a "subalgebra" of $C(S, \mathbb{R})$ or $C(S, \mathbb{C})$ to be dense in the whole space in the uniform metric. Pointwise addition and scalar multiplication make $C(S, \mathbb{R})$ into a real vector space and $C(S, \mathbb{C})$ into a complex vector space, and each space has also the operation of pointwise multiplication; all of these operations on functions preserve continuity as a consequence of Corollary 2.29. By a **subalgebra** of $C(S, \mathbb{R})$ or $C(S, \mathbb{C})$, we mean any nonempty subset that is closed under *all* these operations. The space $C(S, \mathbb{C})$ has also the operation of complex conjugation; this again preserves continuity by Corollary 2.29.

We shall work with a subalgebra of $C(S, \mathbb{R})$ or of $C(S, \mathbb{C})$, and we shall assume that the subalgebra is closed under complex conjugation in the case of complex scalars. The closure of such a subalgebra in the uniform metric is again a subalgebra. To see that this closure is a subalgebra requires checking each operation separately, and we confine our attention to pointwise multiplication. If sequences $\{f_n\}$ and $\{g_n\}$ converge uniformly to f and g, then $\{f_n g_n\}$ converges uniformly to fg because

$$\begin{aligned}
&\sup_{x\in S} |f_n(x)g_n(x) - f(x)g(x)| \\
&\quad\le \sup_{x\in S} |f_n(x)(g_n(x) - g(x))| + \sup_{x\in S} |(f_n(x) - f(x))g(x)| \\
&\quad\le \Big(\sup_{x\in S} |f_n(x)|\Big)\Big(\sup_{x\in S} |g_n(x) - g(x)|\Big) + \Big(\sup_{x\in S} |g(x)|\Big)\Big(\sup_{x\in S} |f_n(x) - f(x)|\Big)
\end{aligned}$$

with $\sup_{x\in S} |g(x)|$ finite and $\sup_{x\in S} |f_n(x)|$ convergent to $\sup_{x\in S} |f(x)|$.

We say that a subalgebra of $C(S, \mathbb{R})$ or $C(S, \mathbb{C})$ **separates points** if for each pair of distinct points x_1 and x_2 in S, there is some f in the subalgebra with $f(x_1) \ne f(x_2)$.

Theorem 2.58 (Stone–Weierstrass Theorem). Let (S, d) be a compact metric space.

(a) If $\mathcal{A}$ is a subalgebra of $C(S, \mathbb{R})$ that separates points and contains the constant functions, then $\mathcal{A}$ is dense in $C(S, \mathbb{R})$ in the uniform metric.

(b) If $\mathcal{A}$ is a subalgebra of $C(S, \mathbb{C})$ that separates points, contains the constant functions, and is closed under complex conjugation, then $\mathcal{A}$ is dense in $C(S, \mathbb{C})$ in the uniform metric.

PROOF OF (a). Let $\mathcal{A}^{\text{cl}}$ be the closure of $\mathcal{A}$ in the uniform metric. We recalled above from Chapter I that $|t|$ is the limit of polynomials $t \mapsto P_n(t)$ uniformly on $[-1, 1]$. It follows that $|t|$ is the limit of polynomials $t \mapsto Q_n(t) = MP_n(M^{-1}t)$ uniformly on $[-M, M]$. Taking $M = \sup_{x \in S} |f(x)|$, we see that $|f|$ is in $\mathcal{A}^{\text{cl}}$ whenever f is in $\mathcal{A}$.

Since $\mathcal{A}^{\text{cl}}$ is a subalgebra closed under addition and scalar multiplication as well, the formulas

$$\max\{f, g\} = \tfrac{1}{2}(f + g) + \tfrac{1}{2}|f - g|,$$
$$\min\{f, g\} = \tfrac{1}{2}(f + g) - \tfrac{1}{2}|f - g|,$$

show that $\mathcal{A}^{\text{cl}}$ is closed under pointwise maximum and pointwise minimum for two functions. Iterating, we see that $\mathcal{A}^{\text{cl}}$ is closed under pointwise maximum and pointwise minimum for n functions for any integer $n \geq 2$.

The heart of the proof is an argument that if $f \in C(S, \mathbb{R})$, $x \in S$, and $\epsilon > 0$ are given, then there exists g_x in $\mathcal{A}^{\text{cl}}$ such that $g_x(x) = f(x)$ and

$$g_x(s) > f(s) - \epsilon$$

for all $s \in S$. The argument is as follows: For each $y \in S$ other than x, there exists a function in $\mathcal{A}$ taking distinct values at x and y. Some linear combination of this function and the constant function 1 is a function h_y in $\mathcal{A}$ with $h_y(x) = f(x)$ and $h_y(y) = f(y)$. To complete the definition of h_y for all $y \in S$, we set h_x equal to the constant function $f(x)1$. The continuity of h_y and the equality $h_y(y) = f(y)$ imply that there exists an open neighborhood U_y of y such that $h_y(s) > f(s) - \epsilon$ for all $s \in U_y$. As y varies, these open neighborhoods cover S, and by compactness of S, finitely many suffice, say $U_{y_1}, \dots U_{y_k}$. Then the function $g_x = \max\{h_{y_1}, \dots, h_{y_k}\}$ has $g_x(s) > f(s) - \epsilon$ for all $s \in S$. Also, it has $g_x(x) = f(x)$, and it is in $\mathcal{A}^{\text{cl}}$, since $\mathcal{A}^{\text{cl}}$ is closed under pointwise maxima.

To complete the proof of (a), we continue with $f \in C(S, \mathbb{R})$ and $\epsilon > 0$ as above. We shall produce a member h of $\mathcal{A}^{\text{cl}}$ such that $|h(s) - f(s)| < \epsilon$ for all $s \in S$. For each x, the continuity of g_x and the equality $g_x(x) = f(x)$ imply that there is an open neighborhood V_x of x such that $g_x(s) < f(s) + \epsilon$ for all $s \in V_x$. As x varies, these open neighborhoods cover S, and by compactness of S, finitely many suffice, say $V_{x_1}, \dots V_{x_l}$. The function $h = \min\{g_{x_1}, \dots, g_{x_l}\}$ has $h(s) < f(s) + \epsilon$ for all $s \in S$, and it is in $(\mathcal{A}^{\text{cl}})^{\text{cl}} = \mathcal{A}^{\text{cl}}$, since each g_{x_j} is in $\mathcal{A}^{\text{cl}}$. Since each g_{x_j} has $g_{x_j}(s) > f(s) - \epsilon$ for all $s \in S$, we have $h(s) > f(s) - \epsilon$ as well. Thus $|h(s) - f(s)| < \epsilon$ for all $s \in S$.

Since ϵ is arbitrary, we conclude that f is a limit point of $\mathcal{A}^{\text{cl}}$. But $\mathcal{A}^{\text{cl}}$ is closed, and hence f is in $\mathcal{A}^{\text{cl}}$. Therefore $\mathcal{A}^{\text{cl}} = C(S, \mathbb{R})$.

PROOF OF (b). Let $\mathcal{A}_{\mathbb{R}}$ be the subset of members of $\mathcal{A}$ that take values in $\mathbb{R}$. Then $\mathcal{A}_{\mathbb{R}}$ is certainly closed under addition, multiplication by real scalars, and pointwise multiplication, and the real-valued constant functions are in $\mathcal{A}_{\mathbb{R}}$. If $f = u + iv$ is in $\mathcal{A}$ and has real and imaginary parts u and v, then $\bar{f}$ is in $\mathcal{A}$ by assumption, and hence so are $u = \frac{1}{2}(f + \bar{f})$ and $v = \frac{1}{2i}(f - \bar{f})$. We are given that $\mathcal{A}$ separates points of S. If x_1 and x_2 are distinct points of S with $f(x_1) \neq f(x_2)$, then either $u(x_1) \neq u(x_2)$ or $v(x_1) \neq v(x_2)$, and it follows that $\mathcal{A}_{\mathbb{R}}$ separates points. By (a), $\mathcal{A}_{\mathbb{R}}$ is dense in $C(S, \mathbb{R})$. Finally let $f = u + iv$ be in $C(S, \mathbb{C})$, and let $\{u_n\}$ and $\{v_n\}$ be sequences in $\mathcal{A}_{\mathbb{R}}$ converging uniformly to u and v, respectively. Then $\{u_n + iv_n\}$ is a sequence in $\mathcal{A}$ converging uniformly to f. Hence $\mathcal{A}$ is dense in $\mathcal{A}_{\mathbb{R}}$.

EXAMPLES.

(1) On a closed bounded interval $[a, b]$ of the line, the scalar-valued polynomials form an algebra that separates points, contains the constants, and is closed under conjugation. The Stone–Weierstrass Theorem in this case reduces to the Weierstrass Theorem (Theorem 1.52), saying that the polynomials are dense in $C([a, b])$.

(2) Consider the algebra of continuous complex-valued periodic functions on $[-\pi, \pi]$ and the subalgebra of complex-valued trigonometric polynomials $\sum_{n=-N}^{N} c_n e^{inx}$; here N depends on the trigonometric polynomial. Neither the algebra nor the subalgebra separates points, since all functions in question have $f(-\pi) = f(\pi)$. To make the theorem applicable, we consider the domain of these functions to be the unit circle of $\mathbb{C}$, parametrized by e^{ix}; this parametrization is permissible by Corollary 1.45, and continuity is preserved. The Stone–Weierstrass Theorem then applies and gives a new proof that the trigonometric polynomials are dense in the space of complex-valued continuous periodic functions; our earlier proof was constructive, deducing the result as part of Fejér's Theorem (Theorem 1.59).

(3) Let S^{n-1} be the unit sphere $\{x \in \mathbb{R}^n \mid |x| = 1\}$ in $\mathbb{R}^n$. The restrictions to S^{n-1} of all scalar-valued polynomials $P(x_1, \dots, x_n)$ in n variables form a subalgebra of $C(S^{m-1})$ that separates points, contains the constants, and is closed under conjugation. The Stone–Weierstrass Theorem says that this subalgebra is dense in $C(S^{n-1})$.

(4) Let S be the closed unit disk $\{z \mid |z| \leq 1\}$ in $\mathbb{C}$. The set $\mathcal{A}$ of restrictions to S of sums of power series having infinite radius of convergence is a subalgebra of $C(S, \mathbb{C})$ that separates points and contains the constants. However, the continuous function $\bar{z}$ is not in the closure, because it has integral 0 over S with every member of $\mathcal{A}$ and also with uniform limits on S of members of $\mathcal{A}$. This example shows the need for some hypothesis like "closed under complex conjugation" in Theorem 2.58b.

Corollary 2.59. If (S, d) is a compact metric space, then $C(S)$ is separable as a metric space.

PROOF. It is enough to consider $C(S, \mathbb{C})$, since $C(S, \mathbb{R})$ is a metric subspace of $C(S, \mathbb{C})$. Being compact metric, S is separable by Proposition 2.33. Let $\mathcal{B}$ be a countable base of S. The number of pairs (U, V) of members of $\mathcal{B}$ such that $U^{\text{cl}} \subseteq V$ is countable. By Proposition 2.30e, there exists a continuous function $f_{UV} : S \to \mathbb{R}$ such that f_{UV} is 1 on U^{cl} and f_{UV} is 0 on V^c. Let us show that the system of functions f_{UV} separates points of S.

If x_1 and x_2 are given, the $\mathbf{T}_1$ property of S (Proposition 2.30a), when combined with Proposition 2.31, gives us a member V of $\mathcal{B}$ such that x_1 is in V and $V \subseteq \{x_2\}^c$. Since the set V^c is closed and does not contain x_1, the property that S is regular (Proposition 2.30c) gives us disjoint open sets U_1 and V_1 with $x_1 \in U_1$ and $V^c \subseteq V_1$. The latter condition means that $V \supseteq V_1^c$. By Proposition 2.31 let U be a basic open set with $x_1 \in U$ and $U \subseteq U_1$. Then we have $x_1 \in U \subseteq U_1 \subseteq U_1^{\text{cl}} \subseteq V_1^c \subseteq V$ and hence also $x_1 \in U \subseteq U^{\text{cl}} \subseteq V$. The function f_{UV} is therefore 1 on x_1 and 0 on x_2, and the system of functions f_{UV} separates points.

The set of all finite products of functions f_{UV} and the constant function 1 is countable, and so is the set D of linear combinations of all these functions with coefficients of the form $q_1 + iq_2$ with q_1 and q_2 rational. The claim is that this countable set D is dense in $C(S, \mathbb{C})$. The closure of D certainly contains the algebra $\mathcal{A}$ of all complex linear combinations of the function 1 and arbitrary finite products of functions f_{UV}, and $\mathcal{A}$ is closed under complex conjugation. By the Stone–Weierstrass Theorem (Theorem 2.58), $\mathcal{A}^{\text{cl}} = C(S, \mathbb{C})$. Since D^{cl} contains $\mathcal{A}$, we have $C(S, \mathbb{C}) = \mathcal{A}^{\text{cl}} \subseteq (D^{\text{cl}})^{\text{cl}} = D^{\text{cl}}$. In other words, D is dense.

11. Completion

If (X, d) and (Y, ρ) are two metric spaces, an **isometry** of X into Y is a function $\varphi : X \to Y$ that preserves distances: $\rho(\varphi(x_1), \varphi(x_2)) = d(x_1, x_2)$ for all x_1 and x_2 in X. For example, a rotation $(x, y) \mapsto (x \cos\theta - y \sin\theta,\ x \sin\theta + y \cos\theta)$ is an isometry of $\mathbb{R}^2$ with itself. An isometry is necessarily continuous (with $\delta = \epsilon$). However, an isometry need not have the whole range as image. For example, the map $x \mapsto (x, 0)$ of $\mathbb{R}^1$ into $\mathbb{R}^2$ is an isometry that is not onto $\mathbb{R}^2$. In the case that there exists an isometry of X onto Y, we say that X and Y are **isometric**.

Theorem 2.60. If (X, d) is a metric space, then there exist a complete metric space (X^*, Δ) and an isometry $\varphi : X \to X^*$ such that the image of X in X^* is dense.

REMARK. It is observed in Problems 25–26 at the end of the chapter that (X^*, Δ) and $\varphi : X \to X^*$ are essentially unique. The metric space (X^*, Δ) is called a **completion** of (X, d), or sometimes "the" completion because of the essential uniqueness. There is more than one construction of X^*, and the proof below will use a construction by Cauchy sequences that is immediately suggested if X is the set of rationals and X^* is the set of reals.

PROOF. Let Cauchy(X) be the set of all Cauchy sequences in X. Define a relation $\sim$ on X as follows: if $\{p_n\}$ and $\{q_n\}$ are in Cauchy(X), then $\{p_n\} \sim \{q_n\}$ means $\lim d(p_n, q_n) = 0$.

Let us prove that $\sim$ is an equivalence relation. It is reflexive, i.e., has $\{p_n\} \sim \{p_n\}$, because $d(p_n, p_n) = 0$ for all n. It is symmetric, i.e., has the property that $\{p_n\} \sim \{q_n\}$ implies $\{q_n\} \sim \{p_n\}$, because $d(p_n, q_n) = d(q_n, p_n)$. It is transitive, i.e., has the property that $\{p_n\} \sim \{q_n\}$ and $\{q_n\} \sim \{r_n\}$ together imply $\{p_n\} \sim \{r_n\}$, because

$$0 \le d(p_n, r_n) \le d(p_n, q_n) + d(q_n, r_n)$$

and each term on the right side is tending to 0. Thus $\sim$ is an equivalence relation.

Let X^* be the set of equivalence classes. If P and Q are two equivalence classes, we set

$$\Delta(P, Q) = \lim d(p_n, q_n), \tag{$*$}$$

where $\{p_n\}$ is a member of the class P and $\{q_n\}$ is a member of the class Q. We have to prove that the limit in $(*)$ exists in $\mathbb{R}$ and then that the limit is independent of the choice of representatives of P and Q.

For the existence of the limit $(*)$, it is enough to prove that the sequence $\{d(p_n, q_n)\}$ is Cauchy. The triangle inequality gives

$$d(p_n, q_n) \le d(p_n, p_m) + d(p_m, q_m) + d(q_m, q_n)$$

and hence $d(p_n, q_n) - d(p_m, q_m) \le d(p_n, p_m) + d(q_m, q_n)$. Reversing the roles of m and n, we obtain

$$|d(p_n, q_n) - d(p_m, q_m)| \le d(p_n, p_m) + d(q_m, q_n).$$

The two terms on the right side tend to 0, since $\{p_k\}$ and $\{q_k\}$ are Cauchy, and hence $\{d(p_n, q_n)\}$ is Cauchy. Thus the limit $(*)$ exists.

We have also to show that the limit $(*)$ is independent of the choice of representatives. Let $\{p_n\}$ and $\{p_n'\}$ be in P, and let $\{q_n\}$ and $\{q_n'\}$ be in Q. Then

$$d(p_n, q_n) \le d(p_n, p_n') + d(p_n', q_n') + d(q_n', q_n).$$

Since the first and third terms on the right side tend to 0 and the other terms in the inequality have limits, we obtain $\lim_n d(p_n, q_n) \le \lim_n d(p_n', q_n')$. Reversing the roles of the primed and unprimed symbols, we obtain $\lim d(p_n', q_n') \le$

$\lim d(p_n, q_n)$. Therefore $\lim d(p_n, q_n) = \lim d(p'_n, q'_n)$, and $\Delta(P, Q)$ is well defined.

Let us see that (X^*, Δ) is a metric space. Certainly $\Delta(P, P) = 0$ and $\Delta(P, Q) = \Delta(Q, P)$. To prove the triangle inequality

$$\Delta(P, Q) \leq \Delta(P, R) + \Delta(R, Q), \tag{$**$}$$

let $\{p_n\}$ be in P, $\{q_n\}$ be in Q, and $\{r_n\}$ be in R. Since

$$d(p_n, q_n) \leq d(p_n, r_n) + d(r_n, q_n),$$

we obtain $(**)$ by passing to the limit. Finally if two unequal classes P and Q are given, and if $\{p_n\}$ and $\{q_n\}$ are representatives, then $\lim d(p_n, q_n) \neq 0$ by definition of $\sim$. Therefore $\Delta(P, Q) > 0$. Thus (X^*, Δ) is a metric space.

Now we can define the isometry $\varphi : X \to X^*$. If x is in X, then $\varphi(x)$ is the equivalence class of the constant sequence $\{p_n\}$ in which $p_n = x$ for all n. To see that φ is an isometry, let x and y be in X, let $p_n = x$ for all n, and let $q_n = y$ for all n. Then $\Delta(\varphi(x), \varphi(y)) = \lim d(p_n, q_n) = \lim d(x, y) = d(x, y)$, and φ is an isometry.

Let us prove that $\varphi(X)$ is dense in X^*. In fact, if P is in X^* and $\{p_n\}$ is a representative, we show that $\varphi(p_n) \to P$. If $\varphi(p_n) = P$ for all sufficiently large n, then P is in $\varphi(X)$; otherwise this limit relation will exhibit P as a limit point of $\varphi(X)$, and we can conclude that P is in $\varphi(X)^{\text{cl}}$ in any case. In other words, $\varphi(p_n) \to P$ implies that $\varphi(X)$ is dense. To prove that we actually do have $\varphi(p_n) \to P$, let $\epsilon > 0$ be given. Choose N such that $k \geq m \geq N$ implies $d(p_m, p_k) < \epsilon$. Then $\Delta(\varphi(p_m), P) = \lim_k d(p_m, p_k) \leq \epsilon$ for $m \geq N$. Hence $\lim_m \Delta(\varphi(p_m), P) = 0$ as required.

Finally let us prove that X^* is complete by showing directly that any Cauchy sequence $\{P_n\}$ converges. Since $\varphi(X)$ is dense in X^*, we can choose $x_n \in X$ with $\Delta(\varphi(x_n), P_n) < 1/n$. First let us prove that $\{x_n\}$ is Cauchy in X. Let $\epsilon > 0$ be given, and choose N large enough so that $\Delta(P_n, P_{n'}) < \epsilon/3$ when n and n' are $\geq N$. Possibly by taking N still larger, we may assume that $1/N < \epsilon/3$. Then whenever n and n' are $\geq N$, we have

$$\begin{aligned} d(x_n, x_{n'}) &= \Delta(\varphi(x_n), \varphi(x_{n'})) \\ &\leq \Delta(\varphi(x_n), P_n) + \Delta(P_n, P_{n'}) + \Delta(P_{n'}, \varphi(x_{n'})) \\ &\leq \tfrac{1}{n} + \tfrac{\epsilon}{3} + \tfrac{1}{n'} \leq \tfrac{1}{N} + \tfrac{\epsilon}{3} + \tfrac{1}{N} < \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} = \epsilon. \end{aligned}$$

Thus $\{x_n\}$ is Cauchy in X. Let $P \in X^*$ be the equivalence class to which $\{x_r\}$ belongs. We prove completeness by showing that $P_n \to P$. Let $\epsilon > 0$ be given, and choose N large enough so that $r \geq n \geq N$ implies $d(x_n, x_r) < \epsilon/2$. Possibly by taking N still larger, we may assume that $\frac{1}{N} < \frac{\epsilon}{2}$. Then $r \geq n \geq N$ implies

$$\Delta(P_n, P) \leq \Delta(P_n, \varphi(x_n)) + \Delta(\varphi(x_n), P) < \tfrac{1}{n} + \lim_r d(x_n, x_r) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon.$$

Thus $P_n \to P$. Hence every Cauchy sequence in X^* converges, and X^* is complete.

An important application of Theorem 2.60 for algebraic number theory is to the construction of the p-adic numbers, p being prime. The metric space that is completed is the set of rationals with a certain nonstandard metric. This application appears in Problems 27–31 at the end of this chapter.

12. Problems

1. As in Example 9 of Section 1, let S be a nonempty set, fix an integer $n > 0$, and let X be the set of n-tuples of members of S. For n-tuples $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, define $d(x, y) = \#\{j \mid x_j \neq y_j\}$, the number of components in which x and y differ. Prove that d satisfies the triangle inequality, so that (X, d) is a metric space.
2. Prove that a separable metric space is the disjoint union of an open set that is at most countable and a closed set in which every point is a limit point.
3. Give an example of a function $f : [0, 1] \to \mathbb{R}$ for which the graph of f, given by $\big\{(x, f(x)) \bigm| 0 \leq x \leq 1\big\}$, is a closed subset of $\mathbb{R}^2$ and yet f is not continuous.
4. If A is a dense subset of a metric space (X, d) and U is open in X, prove that $U \subseteq (A \cap U)^{\mathrm{cl}}$.
5. Let (X, d) be a metric space, let U be an open set, and let $E_1 \supseteq E_2 \supseteq \cdots$ be a decreasing sequence of closed bounded sets with $\bigcap_{n=1}^{\infty} E_n \subseteq U$.
   (a) For X equal to $\mathbb{R}^n$, show that $E_N \subseteq U$ for some N.
   (b) For X equal to the subspace $\mathbb{Q}$ of rationals in $\mathbb{R}^1$, give an example to show that $E_N \subseteq U$ can fail for every N.
6. Let $F : X \times Y \to Z$ be a function from the product of two metric spaces into a metric space.
   (a) Suppose that $(x, y) \mapsto F(x, y)$ is continuous and that Y is compact. Prove that $F(x, \cdot)$ tends to $F(x_0, \cdot)$ uniformly on Y as x tends to x_0.
   (b) Conversely suppose $\mapsto F(x, y)$ is continuous except possibly at points $(x, y) = (x_0, y)$, and suppose that $F(x, \cdot) \to F(x_0, \cdot)$ uniformly. Prove that F is continuous everywhere.

7. Give an example of a continuous function between two metric spaces that fails to carry some Cauchy sequence to a Cauchy sequence.

8. **(Contraction mapping principle)** Let (X, d) be a complete metric space, let r be a number with $0 \le r < 1$, and let $f : X \to X$ be a **contraction mapping**, i.e., a function such that $d(f(x), f(y)) \le rd(x, y)$ for all x and y in X. Prove that there exists a unique x_0 in X such that $f(x_0) = x_0$.

9. Prove that a countable complete metric space has an isolated point.

10. A metric space (X, d) is called **locally connected** if each point has arbitrarily small open neighborhoods that are connected. Let C be a Cantor set in $[0, 1]$, as described in Section 9, and let $X \subset \mathbb{R}^2$ be the union of the three sets $C \times [0, 1]$, $[0, 1] \times \{0\}$, and $[0, 1] \times \{1\}$. Prove that X is compact and connected but is not locally connected.

Problems 11–13 concern the relationship between connected and pathwise connected. It was observed in Section 8 that pathwise connected implies connected. A metric space is called **locally pathwise connected** if each point has arbitrarily small open neighborhoods that are pathwise connected.

11. Prove that a metric space (X, d) that is connected and locally pathwise connected is pathwise connected.

12. Deduce from the previous problem that for an open subset of $\mathbb{R}^n$, connected implies pathwise connected.

13. Prove that any open subset of $\mathbb{R}^1$ is uniquely the disjoint union of open intervals.

Problems 14–17 concern almost periodic functions. Let $f : \mathbb{R}^1 \to \mathbb{C}$ be a bounded uniformly continuous function. If $\epsilon > 0$, an ϵ **almost period** for f is a number t such that $|f(x+t) - f(x)| \le \epsilon$ for all real x. A subset E of $\mathbb{R}^1$ is called **relatively dense** if there is some $L > 0$ such that any interval of length $\ge L$ contains a member of E. The function f is **Bohr almost periodic** if for every $\epsilon > 0$, its set of ϵ almost periods is relatively dense. The function f is **Bochner almost periodic** if every sequence of translates $\{f_{t_n}\}$, where $f_t(x) = f(x + t)$, has a uniformly convergent subsequence. Any function $x \mapsto e^{icx}$ with c real is an example.

14. As usual, let $B(\mathbb{R}^1, \mathbb{C})$ be the metric space of bounded complex-valued functions on $\mathbb{R}^1$ in the uniform metric. Show that the subspace of bounded uniformly continuous functions is closed, hence complete.

15. Show that a bounded uniformly continuous function $f : \mathbb{R}^1 \to \mathbb{C}$ is Bohr almost periodic if and only if the set $\{f_t \mid t \in \mathbb{R}^1\}$ is totally bounded in $B(\mathbb{R}^1, \mathbb{C})$.

16. Prove that a bounded uniformly continuous function $f : \mathbb{R}^1 \to \mathbb{C}$ is Bohr almost periodic if and only if it is Bochner almost periodic. Thus the names Bohr and Bochner can be dropped.

17. Prove that the set of almost periodic functions on $\mathbb{R}^1$ is an algebra closed under complex conjugation and containing the constants. Prove also that it is closed under uniform limits.

Problems 18–20 concern the special case whose proof precedes that of the Stone–Weierstrass Theorem (Theorem 2.58). In the text in Section 10, this preliminary special case was the function $|x|$ on $[-1, 1]$, and it was handled in two ways—in Section I.8 by the binomial expansion and Abel's Theorem and in Section I.9 as a special case of the Weierstrass Approximation Theorem. The problems in the present group handle an alternative preliminary special case, the function $\sqrt{x}$ on $[0, 1]$. This is just as good because $|x| = \sqrt{x^2}$.

18. **(Dini's Theorem)** Let X be a compact metric space. Suppose that $f_n : X \to \mathbb{R}$ is continuous, that $f_1 \le f_2 \le f_3 \le \cdots$, and that $f(x) = \lim f_n(x)$ is continuous and is nowhere $+\infty$. Use the defining property of compactness to prove that f_n converges to f uniformly on X.

19. Define a sequence of polynomial functions $P_n : [0, 1] \to \mathbb{R}$ by $P_0(x) = 0$ and $P_{n+1}(x) = P_n(x) + \frac{1}{2}(x - P_n(x)^2)$. Prove that $0 = P_0 \le P_1 \le P_2 \le \cdots \le \sqrt{x} \le 1$ and that $\lim_n P_n(x) = \sqrt{x}$ for all x in $[0, 1]$.

20. Combine the previous two problems to prove that $\sqrt{x}$ is the uniform limit of polynomial functions on $[0, 1]$.

Problems 21–24 concern the effect of removing from the Stone–Weierstrass Theorem (Theorem 2.58) the hypothesis that the given algebra contains the constants. Let (S, d) be a compact metric space, and let $\mathcal{A}$ be a subalgebra of $C(S, \mathbb{R})$ that separates points. There can be no pair of points $\{x, y\}$ such that all members of $\mathcal{A}$ vanish at x and y.

21. If for each $s \in S$, there is some member of $\mathcal{A}$ that is nonzero at s, prove in the following way that $\mathcal{A}$ is still dense in $C(S, \mathbb{R})$: Observe that the only place in the proof of Theorem 2.58a that the presence of constant functions is used is in the construction of the function h_y in the third paragraph. Show that a function h_y still exists in $\mathcal{A}^{\text{cl}}$ with $h_y(x) = f(x)$ and $h_y(y) = f(y)$ under the weaker hypothesis that for each $s \in S$, there is some member of $\mathcal{A}$ that is nonzero at s.

22. Suppose that the members of $\mathcal{A}$ all vanish at some s_0 in S. Let $\mathcal{B} = \mathcal{A} + \mathbb{R}1$, so that Theorem 2.58a applies to $\mathcal{B}$. Use the linear function $L : C(S, \mathbb{R}) \to \mathbb{R}$ given by $L(f) = f(s_0)$, together with the fact that $\mathcal{B}^{\text{cl}} = C(S, \mathbb{R})$, to prove that $\mathcal{A}$ is uniformly dense in the subalgebra of all members of $C(S, \mathbb{R})$ that vanish at s_0.

23. Adapt the above arguments to prove corresponding results about the algebra $C(S, \mathbb{C})$ of complex-valued continuous functions.

24. Let $C_0([0, +\infty), \mathbb{R})$ be the algebra of continuous functions from $[0, +\infty)$ into $\mathbb{R}$ that have limit 0 at $+\infty$.
    (a) Prove that the set of all finite linear combinations of functions e^{-nx} for positive integers n is dense in $C_0([0, +\infty), \mathbb{R})$.

(b) Suppose that f is in $C_0([0, +\infty), \mathbb{R})$, that $f(x) = 0$ for $x \geq b$, and that $\int_0^b f(x)e^{-nx}\,dx = 0$ for all integers $n > 0$. Prove that f is the 0 function.

Problems 25–26 concern completions of a metric space. They use the notation of Theorem 2.60. The first problem says that the completion is essentially unique, and the second problem addresses the question of what happens if the original space is already complete; in particular it shows that the completion of the completion is the completion.

25. Suppose that (X, d) is a metric space, that (X_1^*, Δ_1) and (X_2^*, Δ_2) are complete metric spaces, and that $\varphi_1 : X \to X_1^*$ and $\varphi_2 : X \to X_2^*$ are isometries such that $\varphi_1(X)$ is dense in X_1^* and $\varphi_2(X)$ is dense in X_2^*. Prove that there exists a unique isometry ψ of X_1^* onto X_2^* such that $\varphi_2 = \psi \circ \varphi_1$.

26. Prove that a metric space X is complete if and only if $X^* = X$, i.e., if and only if the standard isometry φ of X into its completion X^* is onto.

Problems 27–31 concern the field $\mathbb{Q}_p$ of p-**adic numbers**. The problems assume knowledge of unique factorization for the integers; the last problem in addition assumes knowledge of rings, ideals, and quotient rings. Let $\mathbb{Q}$ be the set of rational numbers with their usual arithmetic, and fix a prime number p. Each nonzero rational number r can be written, via unique factorization of integers, as $r = mp^k/n$ with p not dividing m or n and with k a well-defined integer (positive, negative, or zero). Define $|r|_p = p^{-k}$. For $r = 0$, define $|0|_p = 0$. The function $|\cdot|_p$ plays a role in the relationship between $\mathbb{Q}$ and $\mathbb{Q}_p$ similar to the role played by absolute value in the relationship between $\mathbb{Q}$ and $\mathbb{R}$.

27. Prove that $|\cdot|_p$ on $\mathbb{Q}$ satisfies (i) $|r|_p \geq 0$ with equality if and only if $r = 0$, (ii) $|-r|_p = |r|_p$, (iii) $|rs|_p = |r|_p|s|_p$, and (iv) $|r + s|_p \leq \max\{|r|_p, |s|_p\}$. Property (iv) is called the **ultrametric inequality**.

28. Show that $(\mathbb{Q}, d)$ is a metric space under the definition $d(r, s) = |r - s|_p$.

29. Let $(\mathbb{Q}_p, d)$ be the completion of the metric space $(\mathbb{Q}, d)$. Since $|r|_p$ can be recovered from the metric by $|r|_p = d(r, 0)$, the function $|\cdot|_p$ extends to a continuous function $|\cdot|_p : \mathbb{Q}_p \to \mathbb{R}$.
    (a) Using Proposition 2.47, show that addition, as a function from $\mathbb{Q} \times \mathbb{Q}$ to $\mathbb{Q}_p$, extends to a continuous function from $\mathbb{Q}_p \times \mathbb{Q}_p$ to $\mathbb{Q}_p$. Argue similarly that the operation of passing to the negative, as a function from $\mathbb{Q}$ to $\mathbb{Q}_p$, extends to a continuous function from $\mathbb{Q}_p$ to $\mathbb{Q}_p$. Then prove that $\mathbb{Q}_p$ is an abelian group under addition.
    (b) Show that multiplication, as a function from $\mathbb{Q} \times \mathbb{Q}$ to $\mathbb{Q}_p$, extends to a continuous function from $\mathbb{Q}_p \times \mathbb{Q}_p$ to $\mathbb{Q}_p$. (This part is subtler than (a) because multiplication is not uniformly continuous as a function of two variables.)

(c) Let $\mathbb{Q}^\times = \mathbb{Q} - \{0\}$ and $\mathbb{Q}_p^\times = \mathbb{Q}_p - \{0\}$. Show that the operation of taking the reciprocal, as a function from $\mathbb{Q}^\times$ to $\mathbb{Q}_p^\times$, extends to a continuous function from $\mathbb{Q}_p^\times$ to itself. Then prove that $\mathbb{Q}_p^\times$ is an abelian group under multiplication.

(d) Complete the proof that $\mathbb{Q}_p$ is a field by establishing the distributive law $t(r+s) = tr + ts$ within $\mathbb{Q}_p$.

30. (a) Prove that the subset $\{t \in \mathbb{Q}_p \mid |t|_p \le 1\}$ of $\mathbb{Q}_p$ is totally bounded.

(b) Prove that a subset of $\mathbb{Q}_p$ is compact if and only if it is closed and bounded.

31. Prove that the subset $\mathbb{Z}_p$ of $\mathbb{Q}_p$ with $|x|_p \le 1$ is a commutative ring with identity, that the subset P with $|x|_p \le p^{-1}$ is an ideal in $\mathbb{Z}_p$, and that the quotient $\mathbb{Z}_p/P$ is a field of p elements.

CHAPTER III

Theory of Calculus in Several Real Variables

Abstract. This chapter gives a rigorous treatment of parts of the calculus of several variables.

Sections 1–3 handle the more elementary parts of the differential calculus. Section 1 introduces an operator norm that makes the space of linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ or from $\mathbb{C}^n$ to $\mathbb{C}^m$ into a metric space. Section 2 goes through the definitions and elementary facts about differentiation in several variables in terms of linear transformations and matrices. The chain rule and Taylor's Theorem with integral remainder are two of the results of the section. Section 3 supplements Section 2 in order to allow vector-valued and complex-valued extensions of all the results.

Sections 4–5 are digressions. The material in these sections uses the techniques of the present chapter but is not needed until later. Section 4 develops the exponential function on complex square matrices and establishes its properties; it will be applied in Chapter IV. Section 5 establishes the existence of partitions of unity in Euclidean space; this result will be applied at the end of Section 10.

Section 6 returns to the development in Section 2 and proves two important theorems about differential calculus. The Inverse Function Theorem gives sufficient conditions under which a differentiable function from an open set in $\mathbb{R}^n$ into $\mathbb{R}^n$ has a locally defined differentiable inverse, and the Implicit Function Theorem gives sufficient conditions for the local solvability of m nonlinear equations in $n+m$ variables for m of the variables in terms of the other n. The Inverse Function Theorem is proved on its own, and the Implicit Function Theorem is derived from it.

Sections 7–10 treat Riemann integration in several variables. Elementary properties analogous to those in the one-variable case are in Section 7, a useful necessary and sufficient condition for Riemann integrability is established in Section 8, Fubini's Theorem for interchanging the order of integration is in Section 9, and a preliminary change-of-variables theorem for multiple integrals is in Section 10.

1. Operator Norm

This section works with linear functions from n-dimensional column-vector space to m-dimensional column-vector space. It will have applications within this chapter both when the scalars are real and when the scalars are complex. To be neutral let us therefore write $\mathbb{F}$ for $\mathbb{R}$ or $\mathbb{C}$. Material on the correspondence between linear functions and matrices may be found in Section A7 of the appendix.

Specifically let $L(\mathbb{F}^n, \mathbb{F}^m)$ be the vector space of all linear functions from $\mathbb{F}^n$ into $\mathbb{F}^m$. This space corresponds to the vector space of m-by-n matrices with entries in $\mathbb{F}$, as follows: In the notation in Section A7 of the appendix, we let

$(e_1, \dots, e_n)$ be the standard ordered basis of $\mathbb{F}^n$, and $(u_1, \dots, u_m)$ the standard ordered basis of $\mathbb{F}^m$. We define a dot product in $\mathbb{F}^m$ by

$$(a_1, \dots, a_m) \cdot (b_1, \dots, b_m) = a_1 b_1 + \cdots + a_m b_m$$

with no complex conjugations involved. The correspondence of a linear function T in $L(\mathbb{F}^n, \mathbb{F}^m)$ to a matrix A with entries in $\mathbb{F}$ is then given by $A_{ij} = T(e_j) \cdot u_i$.

Let $|\cdot|$ denote the Euclidean norm on $\mathbb{F}^n$ or $\mathbb{F}^m$, given as in Section II.1 by the square root of the sum of the absolute values squared of the entries. The Euclidean norm makes $\mathbb{F}^n$ and $\mathbb{F}^m$ into metric spaces, the distance between two points being the Euclidean norm of the difference.

Proposition 3.1. If T is a member of the space $L(\mathbb{F}^n, \mathbb{F}^m)$ of linear functions from $\mathbb{F}^n$ to $\mathbb{F}^m$, then there exists a finite M such that $|T(x)| \le M|x|$ for all x in $\mathbb{F}^n$. Consequently T is uniformly continuous on $\mathbb{F}^n$.

PROOF. Each x in $\mathbb{F}^n$ has $x = \sum_{j=1}^n (x \cdot e_j) e_j$, and linearity gives $T(x) = \sum_{j=1}^n (x \cdot e_j) T(e_j)$. Thus

$$|T(x)| = \Big| \sum_{j=1}^n (x \cdot e_j) T(e_j) \Big| \le \sum_{j=1}^n |T(e_j)| |x \cdot e_j|.$$

The expression $x \cdot e_j$ is just the j^{th} entry of x, and hence $|x \cdot e_j| \le |x|$. Therefore $|T(x)| \le \big(\sum_{j=1}^n |T(e_j)| \big) |x|$, and the first conclusion has been proved with $M = \sum_{j=1}^n |T(e_j)|$. Replacing x by $x - y$ gives

$$|T(x) - T(y)| = |T(x - y)| \le M|x - y|,$$

and uniform continuity of T follows with $\delta = \epsilon / M$.

Let T be in $L(\mathbb{F}^n, \mathbb{F}^m)$. Using Proposition 3.1, we define the **operator norm** $\|T\|$ of T to be the nonnegative number

$$\|T\| = \inf_{x \in \mathbb{F}^n} \big\{ M \bigm| |T(x)| \le M|x| \text{ for all } x \in \mathbb{F}^n \big\}.$$

Then $$|T(x)| \le \|T\|\,|x| \qquad \text{for all } x \in \mathbb{F}^n.$$

Since $|T(cx)| = |c||x|$ for any scalar c, the inequality $|T(x)| \le M|x|$ holds for all $x \ne 0$ if and only if it holds for all x with $0 < |x| \le 1$, if and only if it holds for all x with $|x| = 1$. Also, we have $T(0) = 0$. It follows that two other expressions for $\|T\|$ are

$$\|T\| = \sup_{|x| \le 1} |T(x)| = \sup_{|x| = 1} |T(x)|.$$

Proposition 3.2. The operator norm on $L(\mathbb{F}^n, \mathbb{F}^m)$ satisfies

(a) $\|T\| \geq 0$ with equality if and only if $T = 0$,
(b) $\|cT\| = |c|\,\|T\|$ for c in $\mathbb{F}$,
(c) $\|T + S\| \leq \|T\| + \|S\|$,
(d) $\|TS\| \leq \|T\|\|S\|$ if S is in $L(\mathbb{F}^n, \mathbb{F}^m)$ and T is in $L(\mathbb{F}^m, \mathbb{F}^k)$,
(e) $\|1\| = 1$ if $n = m$ and 1 denotes the identity function on $\mathbb{F}^n$.

PROOF. All the properties but (d) are immediate. For (d), we have

$$|(TS)(x)| = |T(S(x))| \leq \|T\|\,|S(x)| \leq \|T\|\|S\|\,|x|.$$

Taking the supremum for $|x| \leq 1$ yields $\|TS\| \leq \|T\|\|S\|$.

Corollary 3.3. The space $L(\mathbb{F}^n, \mathbb{F}^m)$ becomes a metric space when a metric d is defined by $d(T, S) = \|T - S\|$.

PROOF. Conclusion (a) of Proposition 3.2 shows that $d(T, S) \geq 0$ with equality if and only if $T = S$, conclusion (b) shows that $d(T, S) = d(S, T)$, and conclusion (c) yields the triangle inequality because substitution of $T = T' - V'$ and $S = V' - U'$ into (c) yields $d(T', U') \leq d(T', V') + d(V', U')$.

Suppose that $\mathbb{F} = \mathbb{C}$. If the matrix A that corresponds to some T in $L(\mathbb{C}^n, \mathbb{C}^m)$ has real entries, we can regard T as a member of $L(\mathbb{R}^n, \mathbb{R}^m)$, as well as a member of $L(\mathbb{C}^n, \mathbb{C}^m)$. Two different definitions of $\|T\|$ are in force. Let us check that they yield the same value for $\|T\|$.

Proposition 3.4. Let T be in $L(\mathbb{C}^n, \mathbb{C}^m)$, and suppose that the vector $T(e_j)$ lies in $\mathbb{R}^m$ for $1 \leq j \leq n$. Then T carries $\mathbb{R}^n$ into $\mathbb{R}^m$, and $\|T\|$ is consistently defined in the sense that

$$\|T\| = \sup_{x \in \mathbb{R}^n,\ |x| \leq 1} |T(x)| = \sup_{z \in \mathbb{C}^n,\ |z| \leq 1} |T(z)|.$$

PROOF. The first conclusion follows since T is $\mathbb{R}$ linear. For the second conclusion, let $\|T\|_{\mathbb{R}}$ and $\|T\|_{\mathbb{C}}$ be the middle and right expressions, respectively, in the displayed equation above. Certainly we have $\|T\|_{\mathbb{R}} \leq \|T\|_{\mathbb{C}}$. If z is in $\mathbb{C}^n$, write $z = x + iy$ with x and y in $\mathbb{R}^n$. Since $T(x)$ and $T(y)$ are in $\mathbb{R}^n$ and T is $\mathbb{C}$ linear,

$$\begin{aligned} |T(z)|^2 &= |T(x) + iT(y)|^2 = |T(x)|^2 + |T(y)|^2 \\ &\leq (\|T\|_{\mathbb{R}}|x|)^2 + (\|T\|_{\mathbb{R}}|y|^2) = \|T\|_{\mathbb{R}}^2(|x|^2 + |y|^2) = \|T\|_{\mathbb{R}}^2|z|^2. \end{aligned}$$

Hence $|T(z)| \leq \|T\|_{\mathbb{R}}|z|$, and it follows that $\|T\|_{\mathbb{C}} \leq \|T\|_{\mathbb{R}}$. The second conclusion follows.

We shall encounter limits of linear functions in the metric d given in Corollary 3.3, and it is worth knowing just what these limits mean. For this purpose, let T be in $L(\mathbb{F}^n, \mathbb{F}^m)$, and define the **Hilbert–Schmidt norm** of T to be

$$|T| = \Big(\sum_{j=1}^n |T(e_j)|^2\Big)^{1/2}.$$

This quantity has an interpretation in terms of the m-by-n matrix A that is associated to the linear function T by the above formula $A_{ij} = T(e_j)\cdot u_i$. Namely, $|T|$ equals $\big(\sum_{i,j} |A_{ij}|^2\big)^{1/2}$, which is just the Euclidean norm of the matrix A if we think of A as lying in $\mathbb{F}^{nm}$. This correspondence provides the license for using the notation of a Euclidean norm for the Hilbert–Schmidt norm of T. The Hilbert–Schmidt norm has the same three properties as the operator norm that allow us to use it to define a metric:

(i) $|T| \geq 0$ with equality if and only if $T = 0$,
(ii) $|cT| = |c|\,|T|$ for c in $\mathbb{F}$,
(iii) $|T + S| \leq |T| + |S|$.

Let us write $d_2(T, S) = |T - S|$ for the associated metric. Parenthetically we might mention that the analogs of (d) and (e) for the Hilbert–Schmidt norm are

(iv) $|TS| \leq |T|\,|S|$ if S is in $L(\mathbb{F}^n, \mathbb{F}^m)$ and T is in $L(\mathbb{F}^m, \mathbb{F}^k)$,
(v) $|1| = \sqrt{n}$ if $n = m$ and 1 denotes the identity function on $\mathbb{F}^n$.

We shall have no need for these last two properties, and their proofs are left to be done in Problem 1 at the end of the chapter.

Proposition 3.5. The operator norm and Hilbert–Schmidt norm on $L(\mathbb{F}^n, \mathbb{F}^m)$ are related by

$$\|T\| \leq |T| \leq \sqrt{n}\,\|T\|.$$

Consequently the associated metrics are related by

$$d \leq d_2 \leq \sqrt{n}\,d.$$

PROOF. If $|x| \leq 1$, then the triangle inequality and the classical Schwarz inequality of Section A5 give

$$\begin{aligned} |T(x)| &= \Big|\sum_{j=1}^n (x\cdot e_j)T(e_j)\Big| \leq \sum_{j=1}^n |x\cdot e_j|\,|T(e_j)| \\ &\leq \Big(\sum_{j=1}^n |x\cdot e_j|\Big)^{1/2}\Big(\sum_{j=1}^n |T(e_j)|^2\Big)^{1/2} = |x|\Big(\sum_{j=1}^n |T(e_j)|^2\Big)^{1/2} \leq |T|. \end{aligned}$$

Taking the supremum over x yields $\|T\| \le |T|$. In addition,

$$|T|^2 = \sum_{j=1}^{n} |T(e_j)|^2 \le \sum_{j=1}^{n} \|T\|^2 |e_j|^2 = n\|T\|^2,$$

and the second asserted inequality follows.

Proposition 3.5 implies that the identity map between the two metric spaces $(L(\mathbb{F}^n, \mathbb{F}^m), d)$ and $(L(\mathbb{F}^n, \mathbb{F}^m), d_2)$ is uniformly continuous and has a uniformly continuous inverse. Therefore open sets, convergent sequences, and even Cauchy sequences are the same in the two metrics. Briefly said, convergence in the operator norm means entry-by-entry convergence of the associated matrices, and similarly for Cauchy sequences.

2. Nonlinear Functions and Differentiation

We begin a discussion of more general functions between Euclidean spaces by defining the multivariable derivative for such a function and giving conditions for its existence. Let E be an open set in $\mathbb{R}^n$, and let $f : E \to \mathbb{R}^m$ be a function. We can write $f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$, where $f_i(x) = f(x) \cdot u_i$. Then $f(x) = \sum_{i=1}^m f_i(x) u_i$. The functions $f_i : E \to \mathbb{R}$ are called the **components** of f. The associated **partial derivatives** are given by

$$\frac{\partial f_i}{\partial x_j}(x) = \frac{d}{dt} f_i(x + te_j)\Big|_{t=0}.$$

We say that f is **differentiable** at x in E if there is some T in $L(\mathbb{R}^n, \mathbb{R}^m)$ with

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

The linear function T is unique if it exists. In fact, if T_1 and T_2 both serve as T in this limit relation, then we write

$$T_2(h) - T_1(h) = \big(f(x+h) - f(x) - T_1(h)\big) - \big(f(x+h) - f(x) - T_2(h)\big)$$

and find that

$$\begin{aligned} \frac{|T_1(h) - T_2(h)|}{|h|} &\le \frac{|f(x+h) - f(x) - T_1(h)|}{|h|} + \frac{|f(x+h) - f(x) - T_2(h)|}{|h|} \\ &\longrightarrow 0. \end{aligned}$$

If $T_1 \neq T_2$, choose some $v \in \mathbb{R}^n$ with $|v| = 1$ and $T_1(v) \neq T_2(v)$. As a nonzero real parameter t tends to 0, we must have

$$\begin{aligned}
&|T_1(v) - T_2(v)| \\
&= |tv|^{-1}\big|\big(f(x+tv) - f(x) - T_1(tv)\big) - \big(f(x+tv) - f(x) - T_2(tv)\big)\big| \\
&\longrightarrow 0.
\end{aligned}$$

Since t does not appear on the left side but the right side tends to 0, the result is a contradiction. Thus $T_1 = T_2$, and T is unique in the definition of "differentiable."

If T exists, we write $f'(x)$ for it and call $f'(x)$ the **derivative** of f at x. If f is differentiable at every point x in E, then $x \mapsto f'(x)$ defines a function $f' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$. We deal with the differentiability of this function presently.

A differentiable function is necessarily continuous. In fact, differentiability at x implies that $|f(x+h) - f(x) - T(h)| \to 0$ as $h \to 0$. Since T is continuous, $T(h) \to 0$ also. Thus $f(x+h) \to f(x)$, and f is continuous at x.

Proposition 3.6. Let E be an open set of $\mathbb{R}^n$, and let $f : E \to \mathbb{R}^m$ be a function. If $f'(x)$ exists, then $\dfrac{\partial f_i}{\partial x_j}(x)$ exists for all i and j, and

$$\frac{\partial f_i}{\partial x_j}(x) = f'(x)(e_j) \cdot u_i.$$

REMARKS. In other words, if $f'(x)$ exists at some point x, then it has to be the linear function whose matrix is $\Big[\dfrac{\partial f_i}{\partial x_j}(x)\Big]$. This matrix is called the **Jacobian matrix** of f at x.

PROOF. We are given that

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

Dot product with a particular vector is continuous by Proposition 3.1. Take $h = te_j$ with t real in the displayed equation, and form the dot product with u_i. Then we obtain

$$\lim_{t \to 0} \frac{|f_i(x+te_j) - f_i(x) - tf'(x)(e_j) \cdot u_i|}{|t|} = 0.$$

The result follows.

The natural converse to Proposition 3.6 is false: the first partial derivatives of a function may all exist at a point, and it can still happen that f is discontinuous.

If $f'(x)$ exists at all points of the open set E in $\mathbb{R}^n$, then we obtain a function $f' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$, and we have seen that we can regard $L(\mathbb{R}^n, \mathbb{R}^m)$ as a Euclidean space by means of the Hilbert–Schmidt norm. Let us examine what continuity of f' means and then what differentiability of f' means.

Theorem 3.7. Let E be an open set of $\mathbb{R}^n$, and let $f : E \to \mathbb{R}^m$ be a function. If $f'(x)$ exists for all x in E and $x \mapsto f'(x)$ is continuous at some x_0, then $x \mapsto \dfrac{\partial f_i}{\partial x_j}(x)$ is continuous at x_0 for all i and j. Conversely if each $\dfrac{\partial f_i}{\partial x_j}(x)$ exists at every point of E and is continuous at a point x_0, then $f'(x_0)$ exists. If all $\dfrac{\partial f_i}{\partial x_j}$ are continuous on E, then $x \mapsto f'(x)$ is continuous on E.

PROOF OF DIRECT PART. The partial derivative $\dfrac{\partial f_i}{\partial x_j}(x)$ is one of the entries of $f'(x)$, regarded as a matrix, and has to be continuous if $f'(x)$ is continuous.

PROOF OF CONVERSE PART. For the moment, let x be fixed. Regard h as $(h_1, \dots, h_m)$, and for $1 \le j \le n$, put $h^{(j)} = (h_1, \dots, h_j, 0, \dots, 0)$. Define T to be the member of $L(\mathbb{R}^n, \mathbb{R}^m)$ with matrix $\Big[\dfrac{\partial f_i}{\partial x_j}(x)\Big]$. Use of the Mean Value Theorem gives

$$\begin{aligned}
[f(x+h) - f(x)]_i &= \sum_{j=1}^n [f(x+h^{(j)}) - f(x+h^{(j-1)})]_i \\
&= \sum_{j=1}^n h_j \frac{d}{dt} f_i(x + h^{(j-1)} + t h_j e_j)\Big|_{t=t_{ij}} \qquad \text{with } 0 < t_{ij} < 1 \\
&= \sum_{j=1}^n h_j \frac{\partial f_i}{\partial x_j}(x + h^{(j-1)} + t_{ij} h_j e_j) \\
&= \sum_{j=1}^n h_j \frac{\partial f_i}{\partial x_j}(x) + \sum_{j=1}^n h_j \Big[\frac{\partial f_i}{\partial x_j}(x + h^{(j-1)} + t_{ij} h_j e_j) - \frac{\partial f_i}{\partial x_j}(x)\Big]
\end{aligned}$$

and hence

$$\frac{[f(x+h) - f(x) - T(h)]_i}{|h|} = \sum_{j=1}^n \frac{h_j}{|h|} \Big[\frac{\partial f_i}{\partial x_j}(x + h^{(j-1)} + t_{ij} h_j e_j) - \frac{\partial f_i}{\partial x_j}(x)\Big].$$

Consequently

$$\frac{|f(x+h) - f(x) - T(h)|}{|h|} \le \sum_{i=1}^m \sum_{j=1}^n \Big|\frac{\partial f_i}{\partial x_j}(x + h^{(j-1)} + t_{ij} h_j e_j) - \frac{\partial f_i}{\partial x_j}(x)\Big|.$$

Let $\epsilon > 0$ be given, and recall that the partial derivatives are assumed to be continuous at x_0. If $\delta > 0$ is chosen such that $|h| < \delta$ implies

$$\Big|\frac{\partial f_i}{\partial x_j}(x_0 + h) - \frac{\partial f_i}{\partial x_j}(x_0)\Big| < \frac{\epsilon}{mn},$$

then we see that $|h| < \delta$ implies

$$\frac{|f(x_0 + h) - f(x_0) - T(h)|}{|h|} < \epsilon.$$

Thus $f'(x_0)$ exists.

Now assume that all the partial derivatives are continuous on E. Since $L(\mathbb{R}^n, \mathbb{R}^m)$ is identified with $\mathbb{R}^{nm}$, the continuity of the entries $\dfrac{\partial f_i}{\partial x_j}(x)$ of the matrix of $f'(x)$ implies the continuity of $f'(x)$ itself. This completes the proof.

If $x \mapsto f'(x)$ is continuous on E, we say that f is of **class** C^1 on E or is a C^1 **function** on E. Let us iterate the above construction: Suppose that E is open in $\mathbb{R}^n$ and that $f : E \to \mathbb{R}^m$ is of class C^1, so that $x \mapsto f'(x)$ is continuous from E into $L(\mathbb{R}^n, \mathbb{R}^m)$. We introduce second partial derivatives of f and the derivative of f'. Namely, define

$$\frac{\partial^2 f_i}{\partial x_k \partial x_j} = \frac{\partial}{\partial x_k}\Big(\frac{\partial f_i}{\partial x_j}\Big).$$

Since the entries of the matrix of $f'(x)$ are $\dfrac{\partial f_i}{\partial x_j}(x) = f'(x)e_j \cdot u_i$, the expression $\dfrac{\partial^2 f_i}{\partial x_k \partial x_j}$ is the partial derivative with respect to x_k of an entry of the matrix of $f'(x)$. Thus we can say that f is of **class** C^2 from E into $\mathbb{R}^m$ if $f'(x)$ is of class C^1, and so on. We say that f is of **class** C^∞ or is a C^∞ **function** if it is of class C^k for all k. A C^∞ function is also said to be **smooth**. We write $C^k(E)$ and $C^\infty(E)$ for the sets of C^k functions and C^∞ functions on E.

Corollary 3.8. Let E be an open set of $\mathbb{R}^n$, and let $f : E \to \mathbb{R}^m$ be a function. The function f is of class C^k on E if and only if all l^{th}-order partial derivatives of each f_i exist and are continuous on E for $l \leq k$.

This is immediate from Theorem 3.7 and the intervening definitions. The definition of a second partial derivative was given in a careful way that stresses the order in which the partial derivatives are to be computed. Reversing the order of two partial derivatives is a problem involving an interchange of limits. In addressing sufficient conditions for this interchange to be valid, it is enough to consider a function of two variables, since $n-2$ variables will remain fixed when we consider a mixed second partial derivative. The different components of the function do not interfere with each other for these purposes, and thus we may assume that the range is $\mathbb{R}^1$.

Proposition 3.9. Let E be an open set in $\mathbb{R}^2$. Suppose that $f : E \to \mathbb{R}^1$ is a function such that $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$, and $\dfrac{\partial^2 f}{\partial y \partial x}$ exist in E and $\dfrac{\partial^2 f}{\partial y \partial x}$ is continuous at $(x, y) = (a, b)$. Then $\dfrac{\partial^2 f}{\partial x \partial y}(a, b)$ exists and equals $\dfrac{\partial^2 f}{\partial y \partial x}(a, b)$.

PROOF. Put

$$\Delta(h, k) = \frac{f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)}{hk},$$

and let $u(t) = f(t, b+k) - f(t, b)$. The function u is a function of one variable t whose derivative is $\frac{\partial f}{\partial x}(t, b+h) - \frac{\partial f}{\partial x}(t, b)$. Use of the Mean Value Theorem produces ξ between a and $a+h$, as well as η between b and $b+k$, such that

$$\begin{aligned} \Delta(h, k) &= \frac{u(a+h) - u(a)}{hk} = \frac{u'(\xi)}{k} \\ &= \frac{\frac{\partial f}{\partial x}(\xi, b+k) - \frac{\partial f}{\partial x}(\xi, b)}{k} = \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta). \end{aligned} \tag{$*$}$$

Let $\epsilon > 0$ be given. By the assumed continuity of $\partial^2 f / \partial y \partial x$ at (a, b), choose $\delta > 0$ such that $|(h, k)| < \delta$ implies

$$\Big| \frac{\partial^2 f}{\partial y \partial x}(a+h, b+k) - \frac{\partial^2 f}{\partial y \partial x}(a, b) \Big| < \epsilon.$$

Then $(*)$ shows that $|(h, k)| < \delta$ implies

$$\Big| \Delta(h, k) - \frac{\partial^2 f}{\partial y \partial x}(a, b) \Big| < \epsilon.$$

Letting k tend to 0 shows, for $|h| < \delta/2$, that

$$\Big| \frac{\frac{\partial f}{\partial y}(a+h, b) - \frac{\partial f}{\partial y}(a, b)}{h} - \frac{\partial^2 f}{\partial y \partial x}(a, b) \Big| \le \epsilon.$$

Since ϵ is arbitrary, $\dfrac{\partial^2 f}{\partial x \partial y}(a, b)$ exists and equals $\dfrac{\partial^2 f}{\partial y \partial x}(a, b)$.

Now that the order of partial derivatives up through order k can be interchanged arbitrarily in the case of a scalar-valued C^k function, we can introduce the usual notation $\dfrac{\partial^k f}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}$ to indicate the result of differentiating f a total of k times, namely k_1 times with respect to x_1, etc., through k_n times with respect to x_n. Simpler notation will be introduced later to indicate such iterated partial derivatives.

Theorem 3.10 (chain rule). Let E be an open set in $\mathbb{R}^n$, and let $f : E \to \mathbb{R}^m$ be a function differentiable at a point x in E. Suppose that g is a function with range $\mathbb{R}^k$ whose domain contains $f(E)$ and is a neighborhood of $f(x)$. Suppose further that g is differentiable at $f(x)$. Then the composition $g \circ f : E \to \mathbb{R}^k$ is differentiable at x, and $(g \circ f)'(x) = g'(f(x))f'(x)$.

PROOF. With x fixed, define $y = f(x)$, $T = f'(x)$, $S = g'(y)$, and also

$$u(h) = f(x+h) - f(x) - T(h) \quad \text{and} \quad v(k) = g(y+k) - g(y) - S(k).$$

Continuity of f at x and of g at y implies that

$$|u(h)| = \varepsilon(h)|h| \quad \text{and} \quad |v(k)| = \eta(k)|k|$$

with $\varepsilon(h)$ tending to 0 as h tends to 0 and with $\eta(k)$ tending to 0 as k tends to 0.

Given $h \neq 0$, put $k = f(x+h) - f(x)$. Then

$$|k| = |T(h) + u(h)| \leq [\|T\| + \varepsilon(h)]|h| \tag{$*$}$$

and

$$\begin{aligned} g(f(x+h)) - g(f(x)) - (ST)(h) &= g(y+k) - g(y) - S(T(h)) \\ &= v(k) + S(k) - S(T(h)) \\ &= S(k - T(h)) + v(k) \\ &= S(u(h)) + v(k). \end{aligned}$$

Therefore

$$\begin{aligned} |h|^{-1}|g(f(x+h)) - g(f(x)) - (ST)(h)| &\leq \|S\|\,|u(h)|/|h| + |v(k)|/|h| \\ &\leq \|S\|\,\varepsilon(h) + \eta(k)|k|/|h| \\ &\leq \|S\|\,\varepsilon(h) + \eta(k)[\|T\| + \varepsilon(h)], \end{aligned}$$

the last inequality following from the upper bound obtained in $(*)$ for $|k|$. As h tends to 0, k tends to 0, by that same bound. Thus $\varepsilon(h)$ and $\eta(k)$ tend to 0. The theorem follows.

Let us clarify in the context of a simple example how the notation in Theorem 3.10 corresponds to the traditional notation for the chain rule. Let f and g be given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = f\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} \qquad \text{and} \qquad z = g\begin{pmatrix} x \\ y \end{pmatrix} = x^2 - y^2.$$

In traditional notation one of the partial derivatives of the composite function is computed by starting from

$$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} = 2x\cos\theta - 2y\sin\theta$$

and then substituting for x and y in terms of r and θ. In notation closer to that of the theorem, we replace derivatives by Jacobian matrices and obtain

$$\begin{aligned} \begin{pmatrix} \dfrac{\partial(g\circ f)}{\partial r} & \dfrac{\partial(g\circ f)}{\partial\theta} \end{pmatrix} &= \begin{pmatrix} \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{pmatrix}\bigg|_{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = f\left(\begin{smallmatrix} r \\ \theta \end{smallmatrix}\right)} \begin{pmatrix} \dfrac{\partial f_1}{\partial r} & \dfrac{\partial f_1}{\partial\theta} \\ \dfrac{\partial f_2}{\partial r} & \dfrac{\partial f_2}{\partial\theta} \end{pmatrix} \\ &= \begin{pmatrix} 2x & -2y \end{pmatrix}\bigg|_{\substack{x = r\cos\theta, \\ y = r\sin\theta}} \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}. \end{aligned}$$

The formula above for $\partial z/\partial r$ is just the first entry of this matrix equation.

The chain rule in several variables is a much more powerful result than its one-variable prototype, permitting one to handle differentiations when a particular variable occurs in several different ways within a function. For example, consider the rule for differentiating a product in one-variable calculus. The function $x \mapsto f(x)g(x)$ can be regarded as a composition if we recognize that one of the ingredients is the multiplication function from $\mathbb{R}^2$ to $\mathbb{R}^1$. Thus let $u = f(x)$ and $v = g(x)$. If we define $F(x) = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}$ and $G\begin{pmatrix} u \\ v \end{pmatrix} = uv$, then $(G\circ F)(x) = f(x)g(x)$. Theorem 3.10 therefore gives

$$\begin{aligned} \frac{d}{dx}(G\circ F)(x) &= \begin{pmatrix} \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v} \end{pmatrix} \begin{pmatrix} f'(x) \\ g'(x) \end{pmatrix} = \begin{pmatrix} v & u \end{pmatrix}\bigg|_{\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right) = F(x)} \begin{pmatrix} f'(x) \\ g'(x) \end{pmatrix} \\ &= \begin{pmatrix} g(x) & f(x) \end{pmatrix} \begin{pmatrix} f'(x) \\ g'(x) \end{pmatrix} = g(x)f'(x) + f(x)g'(x). \end{aligned}$$

Theorem 3.11 (Taylor's Theorem). Let N be an integer ≥ 0, and let E be an open set in $\mathbb{R}^n$. Suppose that $F : E \to \mathbb{R}^1$ is a function of class C^{N+1} on E and that the line segment from $x = (x_1, \dots, x_n)$ to $x + h$, where $h = (h_1, \dots, h_n)$, lies in E. Then

$$F(x+h) = F(x) + \sum_{K=1}^{N} \sum_{\substack{k_1+\cdots+k_n=K,\\ \text{all } k_j \geq 0}} (k_1! \cdots k_n!)^{-1} \frac{\partial^K F(x)}{\partial x_1^{k_1} \cdots x_n^{k_n}} h_1^{k_1} \cdots h_n^{k_n}$$
$$+ \sum_{\substack{l_1+\cdots+l_n=N+1,\\ \text{all } l_j \geq 0}} \frac{N+1}{l_1! \cdots l_n!} h_1^{l_1} \cdots h_n^{l_n} \int_0^1 (1-s)^N \frac{\partial^{N+1} F(x+sh)}{\partial x_1^{l_1} \cdots x_n^{l_n}} \, ds.$$

PROOF. Define a function f of one variable by $f(t) = F(x + th)$. Taylor's Theorem in one variable (Theorem 1.36) gives

$$f(t) = f(0) + \sum_{K=1}^{N} (K!)^{-1} f^{(K)}(0) \, t^K + \frac{1}{N!} \int_0^t (t-s)^N f^{(N+1)}(s) \, ds,$$

and we put $t = 1$ in this formula. If $g(t) = G(x + th)$, the function g is the composition of $t \mapsto x + th$ followed by G, and the chain rule (Theorem 3.10) allows us to compute its derivative as

$$g'(t) = \begin{pmatrix} \dfrac{\partial G}{\partial x_1} & \cdots & \dfrac{\partial G}{\partial x_n} \end{pmatrix} \Bigg|_{x+th} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_{j=1}^{n} h_j \frac{\partial G}{\partial x_j}(x+th).$$

Taking G equal to any of various iterated partial derivatives of F and doing an easy induction, we obtain

$$f^{(K)}(s) = \sum_{\substack{k_1+\cdots+k_n=K,\\ \text{all } k_j \geq 0}} \binom{K}{k_1, \dots, k_n} h_1^{k_1} \cdots h_n^{k_n} \frac{\partial^K F(x+sh)}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}},$$

where $\binom{K}{k_1,\dots,k_n}$ is the multinomial coefficient $\frac{K!}{(k_1)! \cdots (k_n)!}$. Substitution of this expression into the one-variable expansion with $t = 1$ yields the theorem.

3. Vector-Valued Partial Derivatives and Riemann Integrals

It is useful to extend the results of Section 2 so that they become valid for functions $f : E \to \mathbb{C}^m$, where E is an open set in $\mathbb{R}^n$. Up to the chain rule in Theorem

3.10, these extensions are consequences of what has been proved in Section 2 if we identify $\mathbb{C}^m$ with $\mathbb{R}^{2m}$. Achieving the extensions by this identification is preferable to trying to modify the original proofs because of the use of the Mean Value Theorem in the proofs of Theorem 3.7 and Proposition 3.9.

The chain rule extends in the same fashion, once we specify what kinds of functions are to be involved in the composition. We always want the domain to be a subset of some $\mathbb{R}^l$, and thus in a composition $g \circ f$, we can allow g to have values in some $\mathbb{C}^k$, but we insist as in Theorem 3.10 that f have values in $\mathbb{R}^m$.

Now let us turn our attention to Taylor's Theorem as in Theorem 3.11. The statement of Theorem 3.11 allows $\mathbb{R}^1$ as range but not a general $\mathbb{R}^m$. Thus the above extension procedure is not immediately applicable. However, if we allow the given F to take values in $\mathbb{R}^m$, a vector-valued version of Taylor's Theorem will be valid if we adapt our definitions so that the formula remains true component by component. For this purpose we need to enlarge two definitions—that of partial derivatives of any order and that of 1-dimensional Riemann integration—so that both can operate on vector-valued functions. There is no difficulty in doing so, and we may take it that our definitions have been extended in this way.

In the case of vector-valued partial derivatives, let $f : E \to \mathbb{R}^m$ be given. Then $\dfrac{\partial f}{\partial x_j}$ is now defined without passing to components. The entries of this vector-valued partial derivative are exactly the entries of the j^{th} column of the Jacobian matrix of f. Thus the Jacobian matrix consists of the various vector-valued partial derivatives of f, lined up as the columns of the matrix.

Riemann integration is being extended so that the integrand can have values in $\mathbb{R}^m$ or $\mathbb{C}^m$, rather than just $\mathbb{R}^1$. Among the expected properties of the extended version of the Riemann integral, one inequality needs proof because it involves interactions among the various components of the function, namely

$$\Big| \int_a^b F(t)\,dt \Big| \le \int_a^b |F(t)|\,dt.$$

The Riemann integral on the left side is that of a vector-valued function, while the one on the right side is that of a real-valued function. To prove this inequality, let $(\,\cdot\,,\,\cdot\,)$ be the usual inner product for the range space—the dot product if the range is Euclidean space $\mathbb{R}^m$ or the usual Hermitian inner product as in Section II.1 if the range is complex Euclidean space $\mathbb{C}^m$. If u is any vector in the range space with $|u| = 1$, then linearity gives

$$\Big(\int_a^b F(t)\,dt,\ u \Big) = \int_a^b (F(t), u)\,dt.$$

Hence

$$\Big| \Big(\int_a^b F(t)\,dt,\ u \Big) \Big| = \Big| \int_a^b (F(t), u)\,dt \Big| \le \int_a^b |(F(t), u)|\,dt \le \int_a^b |F(t)|\,dt,$$

the two inequalities following from the known scalar-valued version of our inequality and from the Schwarz inequality. If $\int_a^b F(t)\,dt$ is the 0 vector, then our desired inequality is trivial. Otherwise, we specialize the above computation to $u = \big|\int_a^b F(t)\,dt\big|^{-1}\int_a^b F(t)\,dt$, and we obtain our desired inequality.

4. Exponential of a Matrix

In Chapter IV, we shall make use of the **exponential of a matrix** in connection with ordinary differential equations. If A is an n-by-n complex matrix, then we define

$$\exp A = e^A = \sum_{N=0}^{\infty} \frac{1}{N!}\,A^N.$$

This definition makes sense, according to the following proposition.

Proposition 3.12. For any n-by-n complex matrix A, e^A is given by a convergent series entry by entry. Moreover, the series $X \mapsto e^X$ and every partial derivative of an entry of it is uniformly convergent on any bounded subset of matrix space $(= \mathbb{R}^{2n^2})$, and therefore $X \mapsto e^X$ is a C^∞ function.

REMARK. The proof will be tidier if we use derivatives of n-by-n matrix-valued functions. If F and G are two such functions, the same argument as for the usual product rule shows that $\frac{d}{dt}(F(t)G(t)) = F'(t)G(t) + F(t)G'(t)$.

PROOF. Let us define $\|A\|$ for an n-by-n matrix A to be the operator norm of the member of $L(\mathbb{C}^n, \mathbb{C}^n)$ with matrix A. Fix $M \geq 1$. On the set where $\|A\| \leq M$, we have

$$\Big\| \sum_{N=N_1}^{N_2} \frac{1}{N!}\,A^N \Big\| \leq \sum_{N=N_1}^{N_2} \frac{1}{N!}\,\|A^N\| \leq \sum_{N=N_1}^{N_2} \frac{1}{N!}\,\|A\|^N \leq \sum_{N=N_1}^{N_2} \frac{1}{N!} M^n,$$

and the right side tends to 0 uniformly for $\|A\| \leq M$ as N_1 and N_2 tend to infinity. Hence the series for e^A is uniformly Cauchy in the metric built from the operator norm and therefore, by Proposition 3.5, uniformly Cauchy in the metric built from the Hilbert–Schmidt norm. Uniformly Cauchy in the latter metric means that the series is uniformly Cauchy entry by entry, and hence it is uniformly convergent.

The matrices that are 1 or i in one entry and 0 in all other entries form a $2n^2$-member basis over $\mathbb{R}$ of the n-by-n complex matrices. Call these matrices by the names E_j, $1 \leq j \leq 2n^2$. To compute the partial derivative in the E_j direction of a function $f(A)$, we form $\frac{d}{dt} f(A + tE_j)\big|_{t=0}$. We need to estimate the operator norm of a succession of partial derivatives applied to a term $(N!)^{-1}A^N$ of the exponential series. Thus suppose that we have a product $f_1(A)\cdots f_N(A)$ with each $f_i(A)$ equal to A or to a constant, i.e., a matrix that does not depend on A. The partial derivative in the E_j direction of this product is

$$\sum_{i=1}^{N} f_1(A)\cdots f_{i-1}(A)\Big(\tfrac{d}{dt} f_i(A+tE_j)\big|_{t=0}\Big) f_{i+1}(A)\cdots f_N(A).$$

Thus we get a sum of N terms, each involving a sum of the kind of product we are considering. If we repeat this process for partial derivatives in the directions of $E_{j_1},\dots,E_{j_k}$, we get a sum of N^k terms, each involving a sum of the kind of product we are considering. For a factor E_j, Proposition 3.5 gives $\|E_j\| \le |E_j| = 1 \le M$. For a factor of A, we have $\|A\| \le M$. Thus the operator norm of one such product is $\le M^N$. The operator norm of the sum of all such products for a k^{th}-order partial derivative is therefore $\le N^k M^N$. Taking into account the coefficient $1/(N!)$ for the original A^N, we see that the operator norm of terms N_1 through N_2 of the term-by-term k-times differentiated series is

$$\le \sum_{N=N_1}^{N_2} \frac{N^k M^N}{N!}.$$

We see as a consequence that the term-by-term k-times differentiated series obtained from $\sum (N!)^{-1}A^N$ is uniformly convergent entry by entry. By the complex-valued version of Theorem 1.23, applied recursively to handle k^{th}-order partial derivatives, we conclude that $\exp A$ is of class C^k and that the partial derivatives can be computed term by term. Since k is arbitrary, the proof is complete.

Proposition 3.13. The exponential function for matrices satisfies

(a) $e^X e^Y = e^{X+Y}$ if X and Y commute,
(b) e^X is nonsingular,
(c) $\frac{d}{dt}(e^{tX}) = Xe^{tX}$,
(d) $e^{W^{-1}XW} = W^{-1}e^X W$ if W is nonsingular,
(e) $\det e^X = e^{\operatorname{Tr} X}$, where the trace $\operatorname{Tr} X$ is the sum of the diagonal entries of X.

REMARKS. The conclusion of (a) fails for general X and Y, as one sees by taking $X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $Y = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Relevant properties of the determinant function det that appears in the statement of (e) are summarized in Section A7 of the appendix.

PROOF. The rate of convergence determined in Proposition 3.12 is good enough to justify the manipulations that follow. For (a), we have

$$
\begin{aligned}
e^X e^Y &= \Big(\sum_{r=0}^{\infty} \frac{1}{r!} X^r\Big)\Big(\sum_{s=0}^{\infty} \frac{1}{s!} Y^s\Big) = \sum_{r,s\geq 0} \frac{1}{r!s!} X^r Y^s \\
&= \sum_{N=0}^{\infty} \sum_{k=0}^{N} \frac{X^k Y^{N-k}}{k!(N-k)!} = \sum_{N=0}^{\infty} \frac{1}{N!} \sum_{k=0}^{N} \binom{N}{k} X^k Y^{N-k} \\
&= \sum_{N=0}^{\infty} \frac{1}{N!} (X+Y)^N = e^{X+Y}.
\end{aligned}
$$

Conclusion (b) follows by taking $Y = -X$ in (a) and using $e^0 = 1$. For (c), we have

$$
\begin{aligned}
\frac{d}{dt}\big(e^{tX}\big) &= \frac{d}{dt} \sum_{N=0}^{\infty} \frac{1}{N!} (tX)^N = \sum_{N=0}^{\infty} \frac{d}{dt}\Big[\frac{1}{N!} (tX)^N\Big] \\
&= \sum_{N=0}^{\infty} \frac{N}{N!} t^{N-1} X^N = X \sum_{N=1}^{\infty} \frac{1}{(N-1)!} (tX)^{N-1} = X e^{tX}.
\end{aligned}
$$

Conclusion (d) follows from the computation

$$
e^{W^{-1}XW} = \sum_{N=0}^{\infty} \frac{1}{N!} (W^{-1}XW)^N = \sum_{N=0}^{\infty} \frac{1}{N!} W^{-1} X^N W = W^{-1} e^X W.
$$

For conclusion (e), define a complex-valued function f of one variable by $f(t) = \det e^{tX}$. By (a), we have

$$
\begin{aligned}
f'(t) &= \frac{d}{ds} \det e^{(t+s)X}\Big|_{s=0} = \frac{d}{ds} \det(e^{tX} e^{sX})\Big|_{s=0} = \frac{d}{ds} (\det e^{tX})(\det e^{sX})\Big|_{s=0} \\
&= (\det e^{tX}) \frac{d}{ds} (\det e^{sX})\Big|_{s=0} = f(t) \frac{d}{ds} (\det e^{sX})\Big|_{s=0}.
\end{aligned}
$$

Now $e^{sX} = 1 + sX + \frac{1}{2}s^2X^2 + \cdots = 1 + sX + s^2F(s)$ for some smooth matrix-valued function F with entries F_{ij}. If X has entries X_{ij}, then

$$
\begin{aligned}
\det e^{sX} &= \det \begin{pmatrix} 1 + sX_{11} + s^2F_{11}(s) & sX_{12} + s^2F_{12}(s) & \cdots \\ sX_{21} + s^2F_{21}(s) & 1 + sX_{22} + s^2F_{22}(s) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \\
&= 1 + s \operatorname{Tr} X + s^2 G(s)
\end{aligned}
$$

for some smooth function G. Thus $\frac{d}{ds} (\det e^{sX})\big|_{s=0} = \operatorname{Tr} X$, and we obtain $f'(t) = (\operatorname{Tr} X) f(t)$ for all t. Consequently

$$
\frac{d}{dt}\big(e^{-(\operatorname{Tr} X)t} f(t)\big) = e^{-(\operatorname{Tr} X)t} f'(t) - (\operatorname{Tr} X) e^{-(\operatorname{Tr} X)t} f(t) = 0
$$

for all t, and $e^{-(\operatorname{Tr} X)t} f(t)$ is a constant. The constant is seen to be 1 by putting $t = 0$. Therefore $f(t) = e^{(\operatorname{Tr} X)t}$. Conclusion (e) follows by taking $t = 1$.

5. Partitions of Unity

In Section 10 we shall use a "partition of unity" in proving a change-of-variables formula for multiple integrals. As a general matter in analysis, a partition of unity serves as a tool for localizing analysis problems to a neighborhood of each point. The result we shall use in Section 10 is as follows.

Proposition 3.14. Let K be a compact subset of $\mathbb{R}^n$, and let $\{U_1, \dots, U_k\}$ be a finite open cover of K. Then there exist continuous functions $\varphi_1, \dots, \varphi_k$ on $\mathbb{R}^n$ with values in $[0, 1]$ such that

(a) each φ_i is 0 outside of some compact set contained in U_i,
(b) $\sum_{i=1}^k \varphi_i$ is identically 1 on K.

REMARKS. The system $\{\varphi_1, \dots \varphi_k\}$ is an instance of a "partition of unity." For a general metric space X, a **partition of unity** is a family Φ of continuous functions from X into $[0, 1]$ with sum identically 1 such that for each point x in X, there is a neighborhood of x where only finitely many of the functions are not identically 0. The side condition about neighborhoods ensures that the sum $\sum_{\varphi\in\Phi} \varphi(x)$ has only finitely many nonzero terms at each point and that arbitrary partial sums are well-defined continuous functions on X. If $\mathcal{U}$ is an open cover of X, the partition of unity is said to be **subordinate to the cover** $\mathcal{U}$ if each member of Φ vanishes outside some member of $\mathcal{U}$. Further discussion of partitions of unity beyond the present setting appears in the problems at the end of Chapter X. The use of partitions of unity involving continuous functions tends to be good enough for applications to integration problems, but applications to partial differential equations and smooth manifolds are often aided by partitions of unity involving smooth functions, rather than just continuous functions.[1]

We require a lemma.

Lemma 3.15. In $\mathbb{R}^N$,

(a) if L is a compact set and U is an open set with $L \subseteq U$, then there exists an open set V with V^{cl} compact and $L \subseteq V \subseteq V^{\text{cl}} \subseteq U$,
(b) if K is a compact set and $\{U_1, \dots, U_n\}$ is a finite open cover of K, then there exists an open cover $\{V_1, \dots, V_n\}$ of K such that V_i^{cl} is a compact subset of U_i for each i.

[1]Partitions of unity involving smooth functions play no role in the present volume, but they occur in several places in the companion volume *Advanced Real Analysis*, and their existence is addressed there.

PROOF. In (a), if $L = \varnothing$, we can take $V = \varnothing$. If $L \neq \varnothing$, then the continuous function $x \mapsto D(x, U^c)$ on $\mathbb{R}^N$ is everywhere positive on L since $L \subseteq U$. Corollary 2.39 and the compactness of L show that this function attains a positive minimum c on L. If R is chosen large enough so that $L \subseteq B(R; 0)$ and if we take $V = \left\{x \in U \mid D(x, U^c) > \frac{1}{2}c\right\} \cap B(R; 0)$, then $L \subseteq V$, V^{cl} is compact (being closed and bounded), and $V^{\mathrm{cl}} \subseteq \left\{x \in \mathbb{R}^N \mid D(x, U^c) \geq \frac{1}{2}c\right\} \subseteq U$.

For (b), since $\{U_1, \dots, U_n\}$ is a cover of K, we have $K - (U_2 \cup \cdots \cup U_n) \subseteq U_1$. Part (a) produces an open set V_1 with V_1^{cl} compact such that

$$K - (U_2 \cup \cdots \cup U_n) \subseteq V_1 \subseteq V_1^{\mathrm{cl}} \subseteq U_1.$$

The first inclusion shows that $\{V_1, U_2, \dots, U_n\}$ is an open cover of K. Proceeding inductively, let V_i be an open set with

$$K - (V_1 \cup \cdots \cup V_{i-1} \cup U_{i+1} \cup \cdots \cup U_n) \subseteq V_i \subseteq V_i^{\mathrm{cl}} \subseteq U_i.$$

At each stage, $\{V_1, \dots, V_i, U_{i+1}, \dots, U_n\}$ is an open cover of K, and $V_i^{\mathrm{cl}} \subseteq U_i$. Thus $\{V_1, \dots, V_n\}$ is an open cover of K, and $V_i^{\mathrm{cl}} \subseteq U_i$ for all i.

PROOF OF PROPOSITION 3.14. Apply Lemma 3.15b to produce an open cover $\{W_1, \dots, W_k\}$ of K such that W_i^{cl} is compact and $W_i^{\mathrm{cl}} \subseteq U_i$ for each i. Then apply it a second time to produce an open cover $\{V_1, \dots, V_k\}$ of K such that V_i^{cl} is compact and $V_i^{\mathrm{cl}} \subseteq W_i$ for each i. Proposition 2.30e produces a continuous function $g_i \geq 0$ that is 1 on V_i^{cl} and is 0 off W_i. Then $g = \sum_{i=1}^n g_i$ is continuous and ≥ 0 on $\mathbb{R}^n$ and is > 0 everywhere on K. A second application of Proposition 2.30e produces a continuous function $h \geq 0$ that is 1 on the set where g is 0 and is 0 on K. Then $g + h$ is everywhere positive on $\mathbb{R}^n$, and the functions $\varphi_i = g_i/(g + h)$ have the required properties.

6. Inverse and Implicit Function Theorems

The Inverse Function Theorem and the Implicit Function Theorem are results for working with coordinate systems and for defining functions by means of solving equations. Let us use the latter application as a device for getting at the statements of both the theorems.

In the one-variable situation we are given some equation, such as $x^2 + y^2 = a^2$, and we are to think of solving for y in terms of x, choosing one of the possible y's for each x. For example, one solution is $y = -\sqrt{a^2 - x^2}$, $-a < x < a$; unless some requirement like continuity is imposed, there are infinitely many such solutions. In one-variable calculus the terminology is that this solution is "defined implicitly" by the given equation. In terms of functions, the functions $F(x, y) = x^2 + y^2 - a^2$ and $y = f(x) = -\sqrt{a^2 - x^2}$ are such that $F(x, f(x))$ is identically 0. It is then possible to compute dy/dx for this solution in two ways. Only one of these methods remains within the subject of one-variable calculus, namely to compute the "total differential" of $x^2+y^2-a^2$, however that is defined, and to set the result equal to 0. One obtains $2x\,dx + 2y\,dy = 0$ with x and y playing symmetric roles. The declaration that x is to be an independent variable and y is to be dependent means that we solve for dy/dx, obtaining $dy/dx = -x/y$. The other way is more transparent conceptually but makes use of multivariable calculus: it uses the chain rule in two-variable calculus to compute d/dx of $F(x, f(x))$ as the derivative of a composition, the result being set equal to 0 because $(d/dx)F(x, f(x))$ is the derivative of the 0 function. This second method gives $\dfrac{\partial F}{\partial x} + \dfrac{\partial F}{\partial y} f'(x) = 0$, with the partial derivatives evaluated where $(x, y) = (x, f(x))$. Then we can solve for $f'(x)$ provided $\partial F \big/ \partial y$ is not zero at a point of interest, again obtaining $f'(x) = -x/y$. It is an essential feature of both methods that the answer involves both x and y; the reason is that there is more than one choice of y for some x's, and thus specifying x alone does not determine all possibilities for $f'(x)$.

In the general situation we have m equations in $n + m$ variables. Some n of the variables are regarded as independent, and we think in terms of solving for the other m. An example is

$$\begin{aligned} z^3x + w^2y^3 + 2xy &= 0, \\ xyzw - 1 &= 0, \end{aligned}$$

with x and y regarded as the independent variables.

The classical method of implicit differentiation, which is a version of the first method above, is again to form "total differentials"

$$\begin{aligned} 2wy^3\,dw + 3z^2x\,dz + (z^3 + 2y)\,dx + (3w^2y^2 + 2x)\,dy &= 0, \\ xyz\,dw + xyw\,dz + yzw\,dx + xzw\,dy &= 0, \end{aligned}$$

and then to solve the resulting system of equations for dw and dz in terms of dx and dy. The system is

$$\begin{pmatrix} 2wy^3 & 3z^2x \\ xyz & xyw \end{pmatrix} \begin{pmatrix} dw \\ dz \end{pmatrix} = \begin{pmatrix} -(z^3 + 2y)\,dx - (3w^2y^2 + 2x)\,dy \\ -(yzw)\,dx - (xzw)\,dy \end{pmatrix},$$

and the solution is of the form

$$\begin{aligned} dw &= \text{coefficient}\; dx + \text{coefficient}\; dy, \\ dz &= \text{coefficient}\; dx + \text{coefficient}\; dy. \end{aligned}$$

Here the coefficients are the various partial derivatives of interest. Specifically

$$dw = \frac{\partial w}{\partial x}\,dx + \frac{\partial w}{\partial y}\,dy,$$
$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy.$$

The analog of the second method above is to set up matters as a computation of the derivative of a composition. Namely, we write

$$F\begin{pmatrix} x \\ y \\ w \\ z \end{pmatrix} = \begin{pmatrix} z^3x + w^2y^3 + 2xy \\ xyzw - 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} w \\ z \end{pmatrix} = f\begin{pmatrix} x \\ y \end{pmatrix}.$$

We view the given equations as saying that a composition of

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \end{pmatrix}$$

followed by F is the 0 function, i.e.,

$$F\begin{pmatrix} x \\ y \\ f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \end{pmatrix} = 0.$$

We apply the chain rule and compute Jacobian matrices of derivatives, keeping the variables in the same order x, y, w, z. The Jacobian matrix of the 0 function is a 0 matrix of the appropriate size, and the other side of the differentiated equation is the product of two matrices. Thus

$$\begin{aligned}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} z^3 + 2y & 3w^2y^2 + 2x & 2wy^3 & 3z^2x \\ yzw & xzw & xyz & xyw \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{pmatrix} \\ &= \begin{pmatrix} z^3 + 2y & 3w^2y^2 + 2x \\ yzw & xzw \end{pmatrix} + \begin{pmatrix} 2wy^3 & 3z^2x \\ xyz & xyw \end{pmatrix}\begin{pmatrix} \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{pmatrix}.\end{aligned}$$

In other words,

$$\begin{pmatrix} 2wy^3 & 3z^2x \\ xyz & xyw \end{pmatrix}\begin{pmatrix} \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{pmatrix} = -\begin{pmatrix} z^3 + 2y & 3w^2y^2 + 2x \\ yzw & xzw \end{pmatrix},$$

and we have the same system of linear equations as before. Comparing the two methods, we see that we have computed the same things in both methods, merely giving them different names; thus the two methods will lead to the same result in general, not merely in this one example.

The theoretical question is whether the given system of equations, which was $F(x, y, w, z) = 0$ above, can in principle be solved to give a differentiable function; the latter was $\binom{w}{x} = f\begin{pmatrix} x \\ y \end{pmatrix}$ above. The two computational methods show what the partial derivatives are if the equations can be solved, but these methods by themselves give no information about the theoretical question. The theoretical question is answered by the Implicit Function Theorem, which says that there is no problem if the coefficient matrix of our system of linear equations, namely $\begin{pmatrix} 2wy^3 & 3z^2x \\ xyz & xyw \end{pmatrix}$ in the above example, is invertible at a point of interest.

Theorem 3.16 (Implicit Function Theorem). Suppose that F is a C^1 function from an open set E in $\mathbb{R}^{n+m}$ into $\mathbb{R}^m$ and that $F(a, b) = 0$ for some (a, b) in E, with a understood to be in $\mathbb{R}^n$ and b understood to be in $\mathbb{R}^m$. If the matrix $\left[\dfrac{\partial F_i}{\partial y_j}\right]\Big|_{x=a,\, y=b}$ is invertible, then there exist open sets $U \subseteq \mathbb{R}^{n+m}$ and $W \subseteq \mathbb{R}^n$ with (a, b) in U and a in W with this property: to each x in W corresponds a unique y in $\mathbb{R}^m$ such that (x, y) is in U and $F(x, y) = 0$. If this y is defined as $f(x)$, then f is a C^1 function from W into $\mathbb{R}^m$ such that $f(a) = b$, the expression $F(x, f(x))$ is identically 0 for x in W, and

$$f'(x) = -\left[\frac{\partial F_i}{\partial y_j}\right]^{-1}\left[\frac{\partial F_i}{\partial x_j}\right] \qquad \text{at } (x, y) = (x, f(x)).$$

We shall come to the proof shortly. In the example above, $f'(x)$ is the matrix $\begin{pmatrix} \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{pmatrix}$, $\left[\frac{\partial F_i}{\partial y_j}\right]$ is $\begin{pmatrix} 2wy^3 & 3z^2x \\ xyz & xyw \end{pmatrix}$, and $\left[\frac{\partial F_i}{\partial x_j}\right]$ is $\begin{pmatrix} z^3+2y & 3w^2y^2+2x \\ yzw & xzw \end{pmatrix}$.

Let us use the same approach to the question of introducing a new coordinate system in place of an old one. For example, we start with ordinary Euclidean coordinates (u, v) for $\mathbb{R}^2$, and we want to know whether polar coordinates (r, θ) define a legitimate coordinate system in their place. The formula for passing from one system to the other is $\binom{u}{v} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$, but this formula does not really define r and θ. Defining r and θ entails solving for r and θ in terms of u and v. Thus let us set up the system

$$\begin{aligned} r\cos\theta - u &= 0, \\ r\sin\theta - v &= 0. \end{aligned}$$

This is a system of the kind in the Implicit Function Theorem, and the considerations in that theorem apply. The independent vector variable is to be

$x = \binom{u}{v}$, and the dependent vector variable is to be $y = \binom{r}{\theta}$. The system itself is $F(u, v, r, \theta) = 0$, where

$$F(u, v, r, \theta) = \begin{pmatrix} F_1(u, v, r, \theta) \\ F_2(u, v, r, \theta) \end{pmatrix} = \begin{pmatrix} r\cos\theta - u \\ r\sin\theta - v \end{pmatrix}.$$

The sufficient condition for solving the equations locally is that the matrix $\left[\frac{\partial F_i}{\partial y_j}\right]$ be invertible at a point of interest. This is just the matrix

$$\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

The determinant is r, and hence the matrix is invertible except where $r = 0$. The Implicit Function Theorem is therefore telling us in this special case that r and θ give us good *local* coordinates for $\mathbb{R}^2$ except possibly where $r = 0$. The Implicit Function Theorem gives no information about what happens when $r = 0$.

The general result about introducing a new coordinate system in place of an old one is as follows.

Theorem 3.17 (Inverse Function Theorem). Suppose that φ is a C^1 function from an open set E of $\mathbb{R}^n$ into $\mathbb{R}^n$, and suppose that $\varphi'(a)$ is invertible for some a in E. Put $b = \varphi(a)$. Then

(a) there exist open sets $U \subseteq E \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^n$ such that a is in U, b is in V, φ is one-one from U onto V, and
(b) the inverse $f : V \to U$ is of class C^1.

Consequently, $f'(\varphi(x)) = \varphi'(x)^{-1}$ for x in U.

REMARKS. Theorems 3.16 and 3.17 are closely related. We saw in the context of polar coordinates that the Implicit Function Theorem implies the Inverse Function Theorem, and Problem 6 at the end of the chapter points out that this implication is valid in complete generality. Actually, the implication goes both ways, and within this section we shall follow the more standard approach of deriving the Implicit Function Theorem from the Inverse Function Theorem and subsequently proving the Inverse Function Theorem on its own.

PROOF OF THEOREM 3.16 IF THEOREM 3.17 IS KNOWN. Let n, m, E, F, and (a, b) be given as in the statement of Theorem 3.16. We define a function to which we shall apply Theorem 3.17 in dimension $n + m$. The function is

$$\varphi(x, y) = (x, F(x, y)) \quad \text{for } (x, y) \text{ in } E.$$

This satisfies $\varphi(a, b) = (a, F(a, b)) = (a, 0)$, and its Jacobian matrix at (a, b) is

$$\varphi'(a, b) = \begin{pmatrix} \begin{matrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{matrix} & 0 \\ \left[\dfrac{\partial F_i}{\partial x_j}\right]\Big|_{\substack{x=a,\\ y=b}} & \left[\dfrac{\partial F_i}{\partial y_j}\right]\Big|_{\substack{x=a,\\ y=b}} \end{pmatrix}.$$

Since Theorem 3.16 has assumed that $\left[\dfrac{\partial F_i}{\partial y_j}\right]\Big|_{x=a,\, y=b}$ is invertible, $\varphi'(a, b)$ is invertible. Theorem 3.17 therefore applies to φ and produces an open neighborhood W' of $\varphi(a, b) = (a, 0)$ such that φ^{-1} exists on W' and carries W' to an open set. Let $U = \varphi^{-1}(W')$. Define W to be the open neighborhood $W' \cap (\mathbb{R}^n \times \{0\})$ of a in $\mathbb{R}^n$, and define $f(x)$ for x in W by $(x, f(x)) = \varphi^{-1}(x, 0)$. Then f is of class C^1 on W, and $f(a) = b$ because $(a, f(a)) = \varphi^{-1}(a, 0) = (a, b)$. The identity

$$(x, 0) = \varphi(\varphi^{-1}(x, 0)) = \varphi(x, f(x)) = (x, F(x, f(x)))$$

shows that $F(x, f(x)) = 0$ for x in W. The latter equation and the chain rule (Theorem 3.10) give the formula for $f'(x)$.

Finally we are to see that $y = f(x)$ is the unique y in $\mathbb{R}^m$ for which (x, y) is in U and $F(x, y) = 0$. Thus suppose that x is in W and that y_1 and y_2 are in $\mathbb{R}^m$ with (x, y_1) and (x, y_2) in U and $F(x, y_1) = F(x, y_2) = 0$. Then we have $\varphi(x, y_1) = (x, F(x, y_1)) = (x, 0) = (x, F(x, y_2)) = \varphi(x, y_2)$. Since (x, y_1) and (x, y_2) are in U, we can apply φ^{-1} to this equation and obtain $(x, y_1) = (x, y_2)$. Therefore $y_1 = y_2$. This completes the proof of Theorem 3.16 if Theorem 3.17 is known.

Let us turn our attention to a direct proof of the Inverse Function Theorem (Theorem 3.17). When the dimension n is 1, a nonzero derivative at a point yields monotonicity, and the theorem is greatly simplified; this special case is the subject of Section A3 of the appendix.

For general dimension n, it may be helpful to begin with an outline of the proof. The first step is to show that φ is one-one near the point a in question; this is relatively easy. The hard step is to prove that φ is locally onto some open set; this uses either the compactness of closed balls or else their completeness, and we return to a discussion of this step in a moment. The argument for differentiability of the inverse function depends on the continuity of the inverse function; this dependence was already true in the 1-dimensional case in Section A3 of the appendix. Continuity of the inverse function amounts to the fact that small open neighborhoods of a get carried to open sets, and this is part of the proof that φ is locally onto some open set. Finally the chain rule gives

$(\varphi^{-1})'(x) = \big(\varphi'(\varphi^{-1}(x))\big)^{-1}$, and the continuity of $(\varphi^{-1})'$ follows. Thus φ^{-1} is of class C^1.

In carrying out the hard step, one has a choice of using either the compactness of closed balls or else their completeness. The argument using completeness lends itself to certain infinite-dimensional generalizations that are well beyond the scope of this book. Since the argument using compactness is the easier one, we shall use that.

The first step and the hard step mentioned above will be carried out in three lemmas below. After them we address the continuity and differentiability of the inverse function, and the proof of the Inverse Function Theorem will be complete.

Lemma 3.18. If $L : \mathbb{R}^n \to \mathbb{R}^n$ is a linear function that is invertible, then there exists a real number $m > 0$ such that $|L(y)| \geq m|y|$ for all y in $\mathbb{R}^n$.

REMARK. We shall apply this lemma in Lemma 3.19 with $L = \varphi'(a)$.

PROOF. The linear inverse function $L^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is one-one and onto. Thus if y is given, there exists x with $y = L^{-1}(x)$, and we have $|y| = |L^{-1}(x)| \leq \|L^{-1}\||x| \leq \|L^{-1}\||L(y)|$. The lemma follows with $m = \|L^{-1}\|^{-1}$.

Lemma 3.19. In the notation of Theorem 3.17 and Lemma 3.18, choose $m > 0$ such that $|\varphi'(a)(y)| \geq m|y|$ for all $y \in \mathbb{R}^n$, and choose, by continuity of φ', any $\delta > 0$ for which $x \in B(\delta; a)$ implies $\|\varphi'(x) - \varphi'(a)\| \leq \frac{m}{2\sqrt{n}}$. Then $|\varphi(x') - \varphi(x)| \geq \frac{m}{2\sqrt{n}}\,|x' - x|$ whenever x' and x are both in $B(\delta; a)$.

REMARKS. This proves immediately that φ is one-one on $B(\delta; a)$, and it gives an estimate that will establish that φ^{-1} is continuous, once φ^{-1} is known to exist. It proves also that the linear function $\varphi'(x)$ is invertible for $x \in B(\delta; a)$ because

$$\begin{aligned} m|y| &\leq |\varphi'(a)(y)| \\ &\leq |\varphi'(x)(y)| + |\varphi'(x)(y) - \varphi'(a)(y)| \\ &\leq |\varphi'(x)(y)| + \|\varphi'(x) - \varphi'(a)\|\,|y| \\ &\leq |\varphi'(x)(y)| + \frac{m|y|}{2\sqrt{n}}; \end{aligned}$$

if $\varphi'(x)$ were not invertible, then any nonzero y in the kernel of $\varphi'(x)$ would contradict this chain of inequalities.

PROOF. The line segment from x to x' lies within $B(\delta; a)$. Put $z = x' - x$, write this line segment as $t \mapsto x + tz$ for $0 \leq t \leq 1$, and apply the Mean Value Theorem to each component φ_k of φ to obtain

$$\begin{aligned} \varphi_k(x') - \varphi_k(x) &= \varphi_k(x + tz)\big|_{t=1} - \varphi_k(x + tz)\big|_{t=0} \\ &= \varphi'(x + t_k z)(z) \cdot e_k \qquad \text{with } 0 < t_k < 1 \\ &= \varphi'(a)(z) \cdot e_k + (\varphi'(x + t_k z) - \varphi'(a))(z) \cdot e_k. \end{aligned}$$

Taking the absolute value of both sides allows us to write

$$\begin{aligned}|\varphi(x') - \varphi(x)| &\ge |\varphi_k(x') - \varphi_k(x)| \\ &\ge |\varphi'(a)(z) \cdot e_k| - |(\varphi'(x + t_k z) - \varphi'(a))(z)| \\ &\ge |\varphi'(a)(z) \cdot e_k| - \frac{m}{2\sqrt{n}}\,|x' - x|.\end{aligned}$$

Therefore

$$\begin{aligned}|\varphi(x') - \varphi(x)| &\ge \frac{1}{\sqrt{n}}\,|\varphi'(a)(z)| - \frac{m}{2\sqrt{n}}\,|x' - x| \\ &\ge \frac{m}{\sqrt{n}}\,|x' - x| - \frac{m}{2\sqrt{n}}\,|x' - x| \\ &= \frac{m}{2\sqrt{n}}\,|x' - x|.\end{aligned}$$

Lemma 3.20. With notation as in Lemma 3.19, $\varphi(B(\delta; a))$ is open in $\mathbb{R}^n$.

PROOF. Let $c = m\big/(2\sqrt{n}\,)$ be the constant in the statement of Lemma 3.19. Fix x_0 in $B(\delta; a)$ and let $y_0 = \varphi(x_0)$, so that y_0 is the most general element of $\varphi(B(\delta; a))$. Find $\delta_1 > 0$ such that $B(\delta_1; x_0)^{\mathrm{cl}} \subseteq B(\delta; a)$. It is enough to prove that $B(c\delta_1/2; y_0) \subseteq \varphi(B(\delta; a))$. Even better, we prove that $B(c\delta_1/2; y_0) \subseteq \varphi(B(\delta_1; x_0)^{\mathrm{cl}})$.

Thus let y_1 have $|y_1 - y_0| < c\delta_1/2$. Choose, by compactness of $B(\delta_1; x_0)^{\mathrm{cl}}$, a member $x = x_1$ of $B(\delta_1; x_0)^{\mathrm{cl}}$ for which $|\varphi(x) - y_1|^2$ is minimized. Let us show that x_1 is not on the edge of $B(\delta_1; x_0)^{\mathrm{cl}}$, i.e., that $|x_1 - x_0| < \delta_1$. In fact, if $|x_1 - x_0| = \delta_1$, then Lemma 3.19 gives

$$\begin{aligned}|\varphi(x_1) - y_1| &\ge |\varphi(x_1) - y_0| - |y_1 - y_0| \\ &> |\varphi(x_1) - \varphi(x_0)| - c\delta_1/2 \\ &\ge c|x_1 - x_0| - c\delta_1/2 \\ &= c\delta_1/2 \\ &> |y_1 - y_0| \\ &= |\varphi(x_0) - y_1|,\end{aligned}$$

in contradiction to the fact that $|\varphi(x) - y_1|^2$ is minimized on $B(\delta_1; x_0)^{\mathrm{cl}}$ at $x = x_1$. Thus $|x_1 - x_0| < \delta_1$. In this case the scalar-valued function $(\varphi(x) - y_1)\cdot(\varphi(x) - y_1)$ is minimized at an interior point of $B(\delta_1; x_0)^{\mathrm{cl}}$, and all its partial derivatives must be 0. Therefore $\varphi'(x_1)(z) \cdot (\varphi(x_1) - y_1) = 0$ for all z in $\mathbb{R}^n$. Since the linear function $\varphi'(x_1)$ is onto $\mathbb{R}^n$, we conclude that $\varphi(x_1) - y_1 = 0$, and the lemma follows.

COMPLETION OF PROOF OF THEOREM 3.17. Lemma 3.19 showed that the restriction of φ to $B(\delta; a)$ is one-one, and Lemma 3.20 showed that the image is an open set in $\mathbb{R}^n$. Let $f : \varphi(B(\delta; a)) \to B(\delta; a)$ be the inverse function. To complete the proof of Theorem 3.17, we need to see that f is differentiable on $\varphi(B(\delta; a))$. Fix x in $B(\delta; a)$, and suppose that $x + h$ is in $B(\delta; a)$ with $h \neq 0$. Define y and k by $y = \varphi(x)$ and $y + k = \varphi(x + h)$. Since φ is one-one on $B(\delta; a)$, k is not 0. in fact, Lemma 3.19 gives

$$|k| \geq c|h|, \tag{$*$}$$

where $c = m\big/(2\sqrt{n}\,)$. The definitions give

$$\begin{aligned} f(y+k) - f(y) - \varphi'(x)^{-1}(k) &= (x+h) - x - \varphi'(x)^{-1}(k) \\ &= h - \varphi'(x)^{-1}\big(\varphi(x+h) - \varphi(x)\big) \\ &= -\varphi'(x)^{-1}\big(\varphi(x+h) - \varphi(x) - \varphi'(x)(h)\big). \end{aligned}$$

Combining this identity with $(*)$ gives

$$\frac{|f(y+k) - f(y) - \varphi'(x)^{-1}(k)|}{|k|} \leq \frac{\|\varphi'(x)^{-1}\|}{c}\,\frac{|\varphi(x+h) - \varphi(x) - \varphi'(x)(h)|}{|h|}.$$

If $\epsilon > 0$ is given, choose $\eta > 0$ small enough so that

$$\frac{\|\varphi'(x)^{-1}\|}{c}\,\frac{|\varphi(x+h) - \varphi(x) - \varphi'(x)h|}{|h|} < \epsilon$$

as long as $|h| < \eta$. If $|k| < c\eta$, then $|h| < \eta$ by $(*)$ and hence

$$\frac{|f(y+k) - f(y) - \varphi'(x)^{-1}(k)|}{|k|} < \epsilon.$$

In other words, f is differentiable at y, and $f'(y) = \varphi'(x)^{-1}$.

Suppose that the given function φ in the Inverse Function Theorem is better than a C^1 function. What can be said about the inverse function? The answer is carried by the formula $f'(\varphi(x)) = \varphi'(x)^{-1}$ for the derivative of the inverse function f. This formula implies that the partial derivatives of f are quotients of polynomials in partial derivatives of φ by a nonvanishing polynomial (the determinant) in partial derivatives of φ. Thus the iterated partial derivatives of f can be computed harmlessly in terms of the iterated partial derivatives of φ and this same determinant polynomial. Consequently if φ is of class C^k with $k \geq 1$, then so is f. If φ is smooth, so is f. In the case that φ and f are both smooth, we say that φ is a **diffeomorphism**. Let us summarize these facts in a corollary.

Corollary 3.21. Suppose, for some $k \geq 1$, that φ is a C^k function from an open set E of $\mathbb{R}^n$ into $\mathbb{R}^n$, and suppose that $\varphi'(a)$ is invertible for some a in E. Put $b = \varphi(a)$. Let U and V be open subsets of $\mathbb{R}^n$ as in the Inverse Function Theorem such that a is in U, b is in V, and φ is one-one from U onto V. Then the inverse function $f : V \to U$ is of class C^k. If φ is smooth, then φ is a diffeomorphism of U onto V.

7. Definition and Properties of Riemann Integral

Section I.4 contained a careful but limited development of the Riemann integral in one variable. The present section extends that development to several variables. A certain amount of the theory parallels what happened in one variable, and proofs for that part of the theory can be obtained by adjusting the notation and words of Section I.4 in simple ways. Results of that kind are much of the subject matter of this section.

In later sections we shall take up results having no close analog in Section I.4. The main results of this kind are

(i) a necessary and sufficient condition for a function to be Riemann integrable,
(ii) Fubini's Theorem, concerning the relationship between multiple integrals and iterated integrals in the various possible orders,
(iii) a change-of-variables formula for multiple integrals.

We begin a discussion of these in the next section.

The one-variable theory worked with a bounded function $f : [a, b] \to \mathbb{R}$, with domain a closed bounded interval, and we now work with a bounded function $f : A \to \mathbb{R}$ with domain A a "closed rectangle" in $\mathbb{R}^n$. For this purpose a **closed rectangle** (or "closed geometric rectangle") in $\mathbb{R}^n$ is a bounded set of the form

$$A = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

with $a_j \leq b_j$ for all j. Let us abbreviate $[a_j, b_j]$ as A_j. In geometric terms the sides or faces are assumed parallel to the axes or coordinate hyperplanes. We shall use the notion of **open rectangle** in later sections and chapters, an open rectangle being a similar product of bounded open intervals (a_j, b_j) for $1 \leq j \leq n$. However, in this section the term "rectangle" will always mean closed rectangle.

If P_j is a one-variable partition of A_j, then we can form an n-variable **partition** $P = (P_1, \ldots, P_n)$ of the given rectangle A into **component rectangles** $[c_1, d_1] \times \cdots \times [c_n, d_n]$, where c_j and d_j are consecutive subdivision points of P_j. A typical component rectangle is denoted by R, and its n-dimensional volume $\prod_{j=1}^n (d_j - c_j)$ is denoted by $|R|$. The **mesh** $\mu(P)$ of the partition P is the maximum of the meshes of the one-dimensional partitions P_j, hence the largest length of a side of all component rectangles of P.

Relative to our given function f and a given partition P, define $M_R(f) = \sup_{x\in R} f(x)$ and $m_R(f) = \inf_{x\in R} f(x)$ for each component rectangle R of P. Put

$$U(P, f) = \sum_R M_R(f)\,|R| = \text{upper Riemann sum for } P,$$
$$L(P, f) = \sum_R m_R(f)\,|R| = \text{lower Riemann sum for } P,$$
$$\overline{\int_A} f\,dx = \inf_P U(P, f) = \text{upper Riemann integral of } f,$$
$$\underline{\int_A} f\,dx = \sup_P L(P, f) = \text{lower Riemann integral of } f.$$

We say that f is **Riemann integrable** on A if $\overline{\int_A} f\,dx = \underline{\int_A} f\,dx$, and in this case we write $\int_A f\,dx$ for the common value of these two numbers. We write $\mathcal{R}(A)$ for the set of Riemann integrable functions on A. The following lemma is proved in the same way as Lemma 1.24.

Lemma 3.22. Suppose that $f : A \to \mathbb{R}$ has $m \le f(x) \le M$ for all x in A. Then for any partition P of A,

$$m|A| \le L(P, f) \le U(P, f) \le M|A|,$$
$$m|A| \le \underline{\int_A} f\,dx \le M|A|,$$
$$m|A| \le \overline{\int_A} f\,dx \le M|A|.$$

A **refinement** of a partition P of A is a partition P^* such that every component rectangle for P^* is a subset of a component rectangle for P. If $P = (P_1, \ldots, P_n)$ and $P' = (P_1', \ldots, P_n')$ are two partitions of A, then P and P' have at least one common refinement $P^* = (P_1^*, \ldots, P_n^*)$; specifically, for each j, we can take P_j^* to be a common refinement of P_j and P_j'. Arguing as in Lemma 1.25 and Theorem 1.26, we obtain the following two results. The key to the second one of these is the uniform continuity of any continuous function $f : A \to \mathbb{R}$; for the uniform continuity we appeal to the Heine–Borel Theorem (Corollary 2.37) and Proposition 2.41 in several variables, the corresponding one-variable result being Theorem 1.10.

Lemma 3.23. Let $f : A \to \mathbb{R}$ satisfy $m \le f(x) \le M$ for all x in A. Then

(a) $L(P, f) \le L(P^*, f)$ and $U(P^*, f) \le U(P, f)$ whenever P is a partition of A and P^* is a refinement,

(b) $L(P_1, f) \le U(P_2, f)$ whenever P_1 and P_2 are partitions of A,

(c) $\underline{\int}_A f\,dx \le \overline{\int}_A f\,dx$,

(d) $\overline{\int}_A f\,dx - \underline{\int}_A f\,dx \le (M - m)|A|$,

(e) the function f is Riemann integrable on A if and only if for each $\epsilon > 0$, there exists a partition P of A with $U(P, f) - L(P, f) < \epsilon$.

Theorem 3.24. If $f : A \to \mathbb{R}$ is continuous on A, then f is Riemann integrable on A.

Next we argue as in Proposition 1.30 and Theorem 1.31 to obtain two more generalizations to several variables. The several-variable version of uniform continuity is needed in the proof of Proposition 3.25d.

Proposition 3.25. If f_1 and f_2 are Riemann integrable on A, then

(a) $f_1 + f_2$ is in $\mathcal{R}(A)$ and $\int_A (f_1 + f_2)\,dx = \int_A f_1\,dx + \int_A f_2\,dx$,

(b) cf_1 is in $\mathcal{R}(A)$ and $\int_A cf_1\,dx = c\int_A f_1\,dx$ for any real number c,

(c) $f_1 \le f_2$ on A implies $\int_A f_1\,dx \le \int_A f_2\,dx$,

(d) $m \le f_1 \le M$ on A and $\varphi : [m, M] \to \mathbb{R}$ continuous imply that $\varphi \circ f_1$ is in $\mathcal{R}(A)$,

(e) $|f_1|$ is in $\mathcal{R}(A)$, and $\left|\int_A f_1\,dx\right| \le \int_A |f_1|\,dx$,

(f) f_1^2 and $f_1 f_2$ are in $\mathcal{R}(A)$,

(g) $\sqrt{f_1}$ is in $\mathcal{R}(A)$ if $f_1 \ge 0$ on A.

Theorem 3.26. If $\{f_n\}$ is a sequence of Riemann integrable functions on A and if $\{f_n\}$ converges uniformly to f on A, then f is Riemann integrable on A, and $\lim_n \int_A f_n\,dx = \int_A f\,dx$.

There is also a several-variable version of Theorem 1.35, which says that Riemann integrability can be detected by convergence of Riemann sums as the mesh of the partition gets small. Relative to our standard partition $P = (P_1, \ldots, P_n)$, select a member t_R of each component rectangle R relative to P, and define

$$S(P, \{t_R\}, f) = \sum_R f(t_R)|R|.$$

This is called a **Riemann sum** of f.

Theorem 3.27. If f is Riemann integrable on A, then

$$\lim_{\mu(P)\to 0} S(P, \{t_R\}, f) = \int_A f\,dx.$$

Conversely if f is bounded on A and if there exists a real number r such that for any $\epsilon > 0$, there exists some $\delta > 0$ for which $|S(P, \{t_R\}, f) - r| < \epsilon$ whenever $\mu(P) < \delta$, then f is Riemann integrable on A.

REMARK. The proof of the direct part is more subtle in the several-variable case than in the one-variable case, and we therefore include it. The proof of the converse part closely imitates the proof of the converse part of Theorem 1.35, and we omit that.

PROOF. For the direct part the function f is assumed bounded; suppose $|f(x)| \le M$ on A. Let $\epsilon > 0$ be given. Choose a partition $P^* = (P_1^*, \dots, P_n^*)$ of A with $U(P^*, f) \le \int_A f\,dx + \epsilon$. Fix an integer k such that the number of component intervals of P_j^* is $\le k$ for $1 \le j \le n$. Put

$$\delta_1 = \frac{\epsilon}{Mk \sum_{j=1}^n \prod_{i \ne j} |A_i|},$$

and suppose that $P = (P_1, \dots, P_n)$ is any partition of $A = A_1 \times \cdots \times A_n$ with $\mu(P) \le \delta_1$. For each j with $1 \le j \le n$, we separate the component intervals of P_j into two kinds, the ones in $\mathcal{F}^{(j)}$ being the component intervals of P_j that do not lie completely within a single component interval of P_j^* and the ones in $\mathcal{G}^{(j)}$ being the rest. Similarly we separate the component rectangles of P into two kinds, the ones in $\mathcal{F}$ being the component rectangles that do not lie completely within a single component rectangle of P^* and the ones in $\mathcal{G}$ being the rest.

If $R = R_1 \times \cdots \times R_n$ is a member of $\mathcal{F}$, then R_j is in $\mathcal{F}^{(j)}$ for some j with $1 \le j \le n$; let $j = j(R)$ be the first such index. Let $\mathcal{F}_j$ be the subset of R's in $\mathcal{F}$ with $j(R) = j$, so that $\mathcal{F} = \bigcup_{j=1}^n \mathcal{F}_j$ disjointly. Then we have

$$U(P, f) = \sum_{j=1}^n \sum_{R \in \mathcal{F}_j} M_R(f)\,|R| + \sum_{R \in \mathcal{G}} M_R(f)\,|R|. \qquad (*)$$

For the first term on the right side,

$$\begin{aligned}
\Big| \sum_{j=1}^n \sum_{R \in \mathcal{F}_j} M_R(f)\,|R| \Big| &\le M \sum_{j=1}^n \sum_{R \in \mathcal{F}_j} |R| \\
&= M \sum_{j=1}^n \sum_{R_1 \times \cdots \times R_n \in \mathcal{F}_j} |R_1| \times \cdots \times |R_n| \\
&\le M \sum_{j=1}^n \sum_{R_j \in \mathcal{F}^{(j)}} |R_j| \textstyle\prod_{i \ne j} |A_i|.
\end{aligned}$$

Each member R_j of $\mathcal{F}^{(j)}$ contains some point of the partition P_j^* in its interior, and two distinct R_j's cannot contain the same point. Thus the number of R_j's in $\mathcal{F}^{(j)}$ is $\le k$. Also, $|R_j| \le \mu(P)$. Consequently we have

$$\Big|\sum_{j=1}^{n} \sum_{R\in\mathcal{F}_j} M_R(f)\,|R|\Big| \le Mk\mu(P)\sum_{j=1}^{n} \textstyle\prod_{i\ne j}|A_i| \le Mk\delta_1 \sum_{j=1}^{n}\prod_{i\ne j}|A_i| \le \epsilon.$$

The contribution to $U(P, f)$ of the second term on the right side of $(*)$ is

$$\sum_{R\in\mathcal{G}} M_R(f)|R| = \sum_{R^*}\sum_{R\subseteq R^*} M_{R^*}(f)|R| \le \sum_{R^*} M_{R^*}(f)|R^*| \le U(P^*, f).$$

Thus

$$U(P, f) \le \epsilon + U(P^*, f) \le \int_A f\,dx + 2\epsilon.$$

Similarly we can define δ_2 such that $\mu(P) \le \delta_2$ implies

$$L(P, f) \ge \int_A f\,dx - 2\epsilon.$$

If $\delta = \min\{\delta_1, \delta_2\}$ and $\mu(P) \le \delta$, then

$$\int_A f\,dx - 2\epsilon \le L(P, f) \le S(P, \{t_R\}, f) \le U(P, f) \le \int_A f\,dx + 2\epsilon$$

for any choice of points t_R, and hence $\big|S(P, \{t_R\}, f) - \int_A f\,dx\big| \le 2\epsilon$. This completes the proof of the direct part of the theorem.

Finally we include one simple interchange-of-limits result that is handy in working with integrals involving derivatives.

Proposition 3.28. Let f be a complex-valued C^1 function defined on an open set U in $\mathbb{R}^m$, and let K be a compact subset of U. Then

(a) the convergence of $\frac{1}{h}[f(x + he_j) - f(x)]$ to $\frac{\partial f}{\partial x_j}(x)$, as h tends to 0, is uniform for x in K,

(b) the function $g(x_2, \ldots, x_n) = \int_a^b f(x_1, \ldots, x_n)\,dx_1$ is of class C^1 on the set of all points $y = (x_2, \ldots, x_n)$ for which $[a, b] \times \{y\}$ lies in U, and $\frac{\partial}{\partial x_j}\int_a^b f(x)\,dx_1 = \int_a^b \frac{\partial f}{\partial x_j}(x)\,dx_1$ for $j \ne 1$ as long as the set $[a, b] \times \{(x_2, \ldots, x_n)\}$ lies in U.

PROOF. In (a), we may assume that f is real-valued. The Mean Value Theorem gives

$$\frac{1}{h}\big[f(x + he_j) - f(x)\big] - \frac{\partial f}{\partial x_j}(x) = \frac{\partial f}{\partial x_j}(x + te_j) - \frac{\partial f}{\partial x_j}(x)$$

for some t between 0 and h, and then (a) follows from the uniform continuity of $\partial f/\partial x_j$ on K. Conclusion (b) follows by combining (a) and Theorem 1.31.

As we did in the one-variable case in Sections 3 and I.5, we can extend our results concerning integration in several variables to functions with values in $\mathbb{R}^m$ or $\mathbb{C}^m \cong \mathbb{R}^{2m}$. Integration of a vector-valued function is defined entry by entry, and then all the results from Theorem 3.24 through Proposition 3.28 extend. The one thing that needs separate proof is the inequality $\big|\int_A f_1\,dx\big| \le \int_A |f_1|\,dx$ of Proposition 3.25e, and a proof can be carried out in the same way as at the end of Section 3 in the one-variable case.

8. Riemann Integrable Functions

Let E be a subset of $\mathbb{R}^n$. We say that E is of **measure 0** if for any $\epsilon > 0$, E can be covered by a finite or countably infinite set of closed rectangles in the sense of Section 7 of total volume less than ϵ. It is equivalent to require that E can be covered by a finite or countably infinite set of open rectangles of total volume less than ϵ. In fact, if a system of open rectangles covers E, then the system of closures covers E and has the same total volume; conversely if a system of closed rectangles covers E, then the system of open rectangles with the same centers and with sides expanded by a factor $1+\delta$ covers E as long as $\delta > 0$.

Several properties of sets of measure 0 are evident: a set consisting of one point is of measure 0, a face of a closed rectangle is a set of measure 0, and any subset of a set of measure 0 is of measure 0. Less evident is the fact that the countable union of sets of measure 0 is of measure 0. In fact, if $\epsilon > 0$ is given and if $E_1, E_2, \dots$ are sets of measure 0, find finite or countably infinite systems $\mathcal{R}_j$ of closed rectangles for $j \ge 1$ such that the total volume of the members of $\mathcal{R}_j$ is $< \epsilon/2^n$. Then $\mathcal{R} = \bigcup_j \mathcal{R}_j$ is a system of closed rectangles covering $\bigcup_j E_j$ and having total volume $< \epsilon$.

The goal of this section is to prove the following theorem, which gives a useful necessary and sufficient condition for a function of several variables to be Riemann integrable. The theorem immediately extends from the scalar-valued case as stated to the case that f has values in $\mathbb{R}^m$ or $\mathbb{C}^m$.

Theorem 3.29. Let A be a finite closed rectangle in $\mathbb{R}^n$ of positive volume, and let $f : A \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable if and only if the set

$$B = \big\{x \;\big|\; f \text{ is not continuous at } x\big\}$$

has measure 0.

Theorem 3.29 supplies the reassurance that a finite closed rectangle of positive volume cannot have measure 0. In fact, the function f on A that is 1 at every point with all coordinates rational and is 0 elsewhere is discontinuous everywhere on

A. By inspection every $U(P, f)$ is $|A|$ for this f, and every $L(P, f)$ is 0; thus f is not Riemann integrable. The theorem then implies that A is not of measure 0.

The proof of the theorem will make use of an auxiliary notion, that of "content 0," in order to simplify the process of checking whether a given compact set has measure 0. A subset E of $\mathbb{R}^n$ has **content 0** if for any $\epsilon > 0$, E can be covered by a *finite* set of closed rectangles in the sense of Section 7 of total volume less than ϵ. It is equivalent to require that E can be covered by a finite set of open rectangles of total volume less than ϵ. A set consisting of one point is of content 0, a face of a closed rectangle is a set of content 0, any subset of a set of content 0 is of content 0, and the union of *finitely* many sets of content 0 is of content 0.

Every set of content 0 is certainly of measure 0, but the question of any converse relationship is more subtle. Consider the set E of rationals in $[0, 1]$ as a subset of $\mathbb{R}^1$. Since this set is a countable union of one-point sets, it has measure 0. However, it does not have content 0. In fact, if we were to have $E \subseteq \bigcup_{n=1}^N [a_j, b_j]$ with $\sum_{n=1}^N (b_j - a_j) < \epsilon$, then we would have $E^{\text{cl}} \subseteq \bigcup_{n=1}^N [a_j, b_j]$ by Proposition 2.10, since $\bigcup_{n=1}^N [a_j, b_j]$ is closed. Then E^{cl} would have content 0 and necessarily measure 0. This contradicts the fact observed after the statement of the theorem—that a closed rectangle of positive volume, such as $E^{\text{cl}} = [0, 1]$ in $\mathbb{R}^1$, cannot have measure 0. We conclude that a bounded set of measure 0 need not have content 0.

Lemma 3.30. If E is a compact subset of $\mathbb{R}^n$ of measure 0, then E is of content 0.

PROOF. Let E be of measure 0, and let $\epsilon > 0$ be given. Choose open rectangles E_j with $E \subseteq \bigcup_{j=1}^\infty E_j$ and $\sum_{j=1}^\infty |E_j| < \epsilon$. By compactness, $E \subseteq \bigcup_{j=1}^N E_j$ for some N. Then $\sum_{j=1}^N |E_j| < \epsilon$. Since ϵ is arbitrary, E has content 0.

Recall from Section II.9 the notion of the **oscillation** of a function. For a function $f : A \to \mathbb{R}$, the oscillation at x_0 is given by

$$\operatorname{osc}_f(x_0) = \lim_{\delta \downarrow 0} \sup_{x \in B(\delta; x_0)} |f(x) - f(x_0)|.$$

The oscillation is 0 at x_0 if and only if f is continuous there. Lemma 2.55 tells us that

$$\big\{x \in U \bigm| \operatorname{osc}_g(x) \geq 2\epsilon\big\}^{\text{cl}} \subseteq \big\{x \in U \bigm| \operatorname{osc}_g(x) \geq \tfrac{\epsilon}{2}\big\}$$

for any $\epsilon > 0$.

Lemma 3.31. Let A be a nontrivial closed rectangle in $\mathbb{R}^n$, and let $f : A \to \mathbb{R}$ be a bounded function with $\operatorname{osc}_f(x) < \epsilon$ for all x in A. Then there is a partition P of A with $U(P, f) - L(P, f) \leq 2\epsilon |A|$.

PROOF. For each x_0 in A, there is an open rectangle U_{x_0} centered at x_0 such that $|f(x)-f(x_0)| \le \epsilon$ on $A \cap U_{x_0}^{\mathrm{cl}}$. Then $M_{U_{x_0}^{\mathrm{cl}}}(f) - m_{U_{x_0}^{\mathrm{cl}}}(f) \le 2\epsilon$. These open rectangles cover A. By compactness a finite number of them suffice to cover A. Write $A \subseteq U_{x_1} \cup \cdots \cup U_{x_m}$ accordingly. Let P be the partition of A generated by the endpoints in each coordinate of A and the endpoints of the closed rectangles $U_{x_j}^{\mathrm{cl}}$; we discard endpoints that lie outside A. Each component rectangle R of P then lies completely within some $U_{x_j}^{\mathrm{cl}}$, and we have $M_R(f) - m_R(f) \le 2\epsilon$ for each component rectangle R of P. Therefore

$$U(P,f) - L(P,f) = \sum_R \big(M_R(f) - m_R(f)\big)\,|R| \le 2\epsilon \sum_R |R| = 2\epsilon|A|.$$

PROOF OF THEOREM 3.29. Define $B_\epsilon = \big\{x \bigm| \operatorname{osc}_f(x) \ge \epsilon\big\}$ for each $\epsilon > 0$, so that $B = \bigcup_{n=1}^\infty B_{1/n}$. For the easy direction of the proof, suppose that f is Riemann integrable. We show that $B_{1/n}$ has content 0 for all n. Since content 0 implies measure 0, B_n will have measure 0 for all n. So will the countable union, and therefore B will have measure 0.

Given $\epsilon > 0$ and n, use Lemma 3.23e to choose a partition P of A with $U(P,f) - L(P,f) \le \epsilon/n$. Let

$$\mathcal{R} = \big\{\text{component rectangles } R \text{ of } P \bigm| R^o \cap B_{1/n} \ne \varnothing\big\},$$

where R^o is the interior of R. Then $B_{1/n}$ is covered by the closed rectangles in $\mathcal{R}$ and the boundaries of the component rectangles of P. The latter are of content 0. For R in $\mathcal{R}$, let us see that $M_R(f) - m_R(f) \ge 1/n$. In fact, if x_0 is in $R^o \cap B_{1/n}$, then $\operatorname{osc}_f(x_0) \ge 1/n$, so that $\lim_{\delta \downarrow 0} \sup_{|x-x_0|<\delta} |f(x) - f(x_0)| \ge 1/n$ and

$$\sup_{\substack{|x-x_0|<\delta,\\ x\in R^o}} |f(x) - f(x_0)| \ge 1/n \qquad \text{for all } \delta > 0.$$

Therefore $M_R(f) - m_R(f) \ge 1/n$. Summing on $R \in \mathcal{R}$ gives

$$\begin{aligned}\frac{1}{n}\sum_{R\in\mathcal{R}} |R| \le \sum_{R\in\mathcal{R}} \big(M_R(f) - m_R(f)\big)\,|R| &\le \sum_{\text{all } R} \big(M_R(f) - m_R(f)\big)\,|R| \\ &= U(P,f) - L(P,f) \le \epsilon/n,\end{aligned}$$

and thus $\sum_{R\in\mathcal{R}} |R| \le \epsilon$. Consequently $B_{1/n}$ has content 0, as asserted.

For the converse direction of the proof, suppose that B has measure 0. We are to prove that f is Riemann integrable. Let $\epsilon > 0$ be given. The inclusion of Lemma 2.55 gives $B_\epsilon^{\mathrm{cl}} \subseteq B_{\epsilon/4} \subseteq B$, and thus B_ϵ^{cl} has measure 0. The set B_ϵ^{cl} is compact, and Lemma 3.30 shows that it has content 0. Hence the subset

B_ϵ has content 0. Choose open rectangles $U_1, \ldots, U_m$ such that $B_\epsilon \subseteq \bigcup_{j=1}^m U_j$ and $\sum_{j=1}^m |U_j| < \epsilon$. Form the partition P of A generated by the endpoints in each coordinate of A and the endpoints of the closed rectangles $U_{x_j}^{\text{cl}}$; we discard endpoints that lie outside A.

Then every component closed rectangle R of P is in one of the two classes

$$\mathcal{R}_1 = \{R \mid R \subseteq U_j^{\text{cl}} \text{ for some } j\},$$
$$\mathcal{R}_2 = \{R \mid R \cap B_\epsilon = \varnothing\}.$$

In fact, our definition is such that $R \cap U_j \neq \varnothing$ implies $R \subseteq U_j^{\text{cl}}$. If $R \cap B_\epsilon \neq \varnothing$, let x_0 be in $R \cap B_\epsilon$. Then x_0 is in some U_j, and $R \cap U_j \neq \varnothing$ for that j. Hence R is in $\mathcal{R}_1$.

We shall construct a particular refinement P' of P in a moment. Let R' be a typical component rectangle of P'. For any refinement P' of P, we have

$$\begin{aligned}
&U(P', f) - L(P', f) \\
&\quad \leq \sum_{R \in \mathcal{R}_1} \sum_{R' \subseteq R} \big(M_{R'}(f) - m_{R'}(f)\big)\,|R'| + \sum_{R \in \mathcal{R}_2} \sum_{R' \subseteq R} \big(M_{R'}(f) - m_{R'}(f)\big)\,|R'| \\
&\quad \leq 2\big(\sup_A |f|\big) \sum_{R \in \mathcal{R}_1} \sum_{R' \subseteq R} |R'| + \sum_{R \in \mathcal{R}_2} \sum_{R' \subseteq R} \big(M_{R'}(f) - m_{R'}(f)\big)\,|R'| \\
&\quad \leq 2\big(\sup_A |f|\big)\epsilon + \sum_{R \in \mathcal{R}_2} \sum_{R' \subseteq R} \big(M_{R'}(f) - m_{R'}(f)\big)\,|R'|
\end{aligned}$$

since $\sum_{j=1}^m |U_j| < \epsilon$. For R in $\mathcal{R}_2$, we have $\operatorname{osc}_f(x) < \epsilon$ for all x in R. Lemma 3.31 shows that there is a partition P_R of R such that $U(P_R, f) - L(P_R, f) \leq 2\epsilon|R|$. In other words, $\sum_{R' \subseteq R} \big(M_{R'}(f) - m_{R'}(f)\big)\,|R'| \leq 2\epsilon|R|$ if P' is fine enough to include all the n-tuples of P_R. If P' is fine enough so that this happens for all R in $\mathcal{R}_2$, then we obtain

$$U(P', f) - L(P', f) \leq 2\big(\sup_A |f|\big)\epsilon + \sum_{R \in \mathcal{R}_2} 2\epsilon|R| \leq 2\epsilon\big(\sup_A |f| + |A|\big),$$

and the theorem follows.

9. Fubini's Theorem for the Riemann Integral

Fubini's Theorem is a result asserting that a double integral is equal to an iterated integral in either order. An unfortunate feature of the Riemann integral is that when an integrable function $f(x, y)$ is restricted to one of the two variables, then the resulting function of that variable need not be integrable. Thus a certain amount of checking is often necessary in using the theorem. This feature is corrected in the Lebesgue integral, and that, as we shall see in Chapter V, is one of the strengths of the Lebesgue integral.

Theorem 3.32 (Fubini's Theorem). Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be closed rectangles, and let $f : A \times B \to \mathbb{R}$ be Riemann integrable. For x in A let f_x be the function $y \mapsto f(x, y)$ for y in B, and define

$$\mathcal{L}(x) = \underline{\int}_B f_x(y)\,dy = \underline{\int}_B f(x, y)\,dy,$$
$$\mathcal{U}(x) = \overline{\int}_B f_x(y)\,dy = \overline{\int}_B f(x, y)\,dy,$$

as functions on A. Then $\mathcal{L}$ and $\mathcal{U}$ are Riemann integrable on A and

$$\int_{A\times B} f\,dx\,dy = \int_A \mathcal{L}(x)\,dx = \int_A \Big[\underline{\int}_B f(x, y)\,dy\Big]\,dx,$$
$$\int_{A\times B} f\,dx\,dy = \int_A \mathcal{U}(x)\,dx = \int_A \Big[\overline{\int}_B f(x, y)\,dy\Big]\,dx.$$

PROOF. Let P be a partition of the form (P_A, P_B), and let $R = R_A \times R_B$ be a typical component rectangle of P. Then

$$L(P, f) = \sum_R m_R(f)\,|R| = \sum_{R_A} \Big(\sum_{R_B} m_{R_A\times R_B}(f)\,|R_B| \Big)|R_A|.$$

For x in R_A, $m_{R_A\times R_B}(f) \le m_{R_B}(f_x)$. Hence x in R_A implies

$$\sum_{R_B} m_{R_A\times R_B}(f)\,|R_B| \le \sum_{R_B} m_{R_B}(f_x)\,|R_B| \le \underline{\int}_B f_x\,dy = \mathcal{L}(x).$$

Taking the infimum over x in R_A and summing over R_A gives

$$L(P, f) = \sum_{R_A} \Big(\sum_{R_B} m_{R_A\times R_B}(f)\,|R_B| \Big)|R_A| \le \sum_{R_A} m_{R_A}(\mathcal{L})\,|R_A| = L(P_A, \mathcal{L}).$$

Similarly

$$U(P_A, \mathcal{U}) \le U(P, f).$$

Thus

$$L(P, f) \le L(P_A, \mathcal{L}) \le U(P_A, \mathcal{L}) \le U(P_A, \mathcal{U}) \le U(P, f).$$

Since f is Riemann integrable, the ends of the above display can be made close together by choosing P appropriately. The second and third members of the display will then be close, and hence

$$\int_{A\times B} f\,dx\,dy = \underline{\int}_A \mathcal{L}\,dx = \overline{\int}_A \mathcal{L}\,dx.$$

The result for $\mathcal{L}$ follows. The result for $\mathcal{U}$ follows in similar fashion immediately from the inequalities

$$L(P, f) \le L(P_A, \mathcal{L}) \le L(P_A, \mathcal{U}) \le U(P_A, \mathcal{U}) \le U(P, f).$$

This proves the theorem.

REMARKS.

(1) Equality of the double integral with the iterated integral in the other order is the same theorem. Thus the iterated integrals in the two orders are equal.

(2) If f is continuous on $A \times B$, then f_x is continuous on B as a consequence of Corollary 2.27, so that $\underline{\int}_B f(x, y)\, dy = \overline{\int}_B f(x, y)\, dy$. Hence

$$\int_{A\times B} f\, dx\, dy = \int_A \Big[\int_B f(x, y)\, dy \Big]\, dx$$

when f is continuous on $A \times B$. This result is isolated as Corollary 3.33 below. Evidently it immediately extends to continuous functions with values in $\mathbb{R}^k$ or $\mathbb{C}^k$.

(3) In practice one often considers integrals of the form $\int_U f(x, y)\, dx\, dy$ for some open set U, where f is continuous on some closed rectangle $A \times B$ containing U. Then the double integral equals $\int_{A\times B} f(x, y) I_U(x, y)\, dx\, dy$, where I_U is the **indicator function**[2] of U equal to 1 on U and 0 off U. In many applications the functions $(I_U)_x$ have harmless discontinuities and $(f I_U)_x$ is therefore Riemann integrable as a function of x. In this case, the upper and lower integrals can again be dropped in the statement of Theorem 3.32.

Corollary 3.33 (Fubini's Theorem for continuous integrand). Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be closed rectangles, and let $f : A \times B \to \mathbb{R}$ be continuous. Then

$$\int_{A\times B} f\, dx\, dy = \int_A \Big[\int_B f(x, y)\, dy \Big]\, dx = \int_B \Big[\int_A f(x, y)\, dx \Big]\, dy.$$

10. Change of Variables for the Riemann Integral

The goal in this section is to prove a several-variables generalization of the one-variable formula

$$\int_a^b f(x)\, dx = \int_A^B f(\varphi(y))\varphi'(y)\, dy$$

given in Theorem 1.34. In the one-variable case we assumed in effect that φ was a strictly increasing function of class C^1 on $[a, b]$ and that f was merely Riemann integrable. The several-variables theorem in this section will be only a preliminary result, with a final version stated and proved in Chapter VI in the

[2]Indicator functions are called "characteristic functions" by many authors, but the term "characteristic function" has another meaning in probability theory and is best avoided as a substitute for "indicator function" in any context where probability might play a role.

context of the Lebesgue integral. In particular we shall assume in the present section that f is continuous and that it vanishes near the boundary of the domain, and we shall make strong assumptions about φ. To capture succinctly the notion that f vanishes near the boundary of its domain, we introduce the notion of the **support** of f, which is the closure of the set where f is nonzero.

Theorem 3.34 (change-of-variables formula). Let φ be a one-one function of class C^1 from an open subset U of $\mathbb{R}^n$ onto an open subset $\varphi(U)$ of $\mathbb{R}^n$ such that $\det \varphi'(x)$ is nowhere 0. Then

$$\int_{\varphi(U)} f(y)\,dy = \int_U f(\varphi(x))|\det \varphi'(x)|\,dx$$

for every continuous function $f : \varphi(U) \to \mathbb{R}$ whose support is a compact subset of $\varphi(U)$.

Before a discussion of the sense in which this result has to regarded as preliminary, a few remarks are in order. The function φ' is the usual derivative of φ, and $\varphi'(x)$ is therefore a linear function from $\mathbb{R}^n$ to $\mathbb{R}^n$ that depends on x. The matrix of the linear function $\varphi'(x)$ is the Jacobian matrix $[\partial\varphi_i/\partial x_j]$, and $\det \varphi'(x)$ is the determinant of this matrix. In classical notation, this determinant is often written as $\dfrac{\partial(y_1,\dots,y_n)}{\partial(x_1,\dots,x_n)}$, and then the effect on the integral of changing variables can be summarized by the formula $dy = \left|\dfrac{\partial(y_1,\dots,y_n)}{\partial(x_1,\dots,x_n)}\right| dx$. The absolute value signs did not appear in the one-variable formula in Theorem 1.34, but the assumption that φ was strictly increasing made them unnecessary, $\varphi'(x)$ being > 0. Had we worked with strictly decreasing φ, we would have assumed $\varphi'(x) < 0$ everywhere, and the limits of integration on one side of the formula would have been reversed from their natural order. The minus sign introduced by putting the limits of integration in their natural order would have compensated for a minus sign introduced in changing $\varphi'(x)$ to $|\varphi'(x)|$.

The hypotheses on φ make the Inverse Function Theorem (Theorem 3.17) applicable at every x in U. Consequently $\varphi(U)$ is automatically open, and φ has a locally defined C^1 inverse function about each point $\varphi(x)$ of the image. Since φ has been assumed to be one-one, $\varphi : U \to \varphi(U)$ has a global inverse function φ^{-1} of class C^1.

We can use φ^{-1} to verify that $f \circ \varphi$ has compact support in U. To the equality $\varphi\big(\{x \in U \mid f(\varphi(x)) \neq 0\}\big) = \{y \in \varphi(U) \mid f(y) \neq 0\}$, we apply φ^{-1} and obtain $\{x \in U \mid f(\varphi(x)) \neq 0\} = \varphi^{-1}\big(\{y \in \varphi(U) \mid f(y) \neq 0\}\big)$. Hence

$$\{x \in U \mid f(\varphi(x)) \neq 0\}^{\mathrm{cl}} = \big(\varphi^{-1}\big(\{y \in \varphi(U) \mid f(y) \neq 0\}\big)\big)^{\mathrm{cl}}.$$

The identity $F(E^{\text{cl}}) \subseteq (F(E))^{\text{cl}}$ holds whenever F is a continuous function between two metric spaces, by Proposition 2.25. When E^{cl} is compact, equality actually holds. The reason is that Propositions 2.34 and 2.38 show $F(E^{\text{cl}})$ to be closed; since $F(E^{\text{cl}})$ is a closed set containing $F(E)$, it contains $(F(E))^{\text{cl}}$. Applying this fact to the displayed equation above, we obtain

$$\big\{x \in U \,\big|\, f(\varphi(x)) \neq 0\big\}^{\text{cl}} = \varphi^{-1}\big(\big\{y \in \varphi(U) \,\big|\, f(y) \neq 0\big\}^{\text{cl}}\big).$$

In other words,

$$\operatorname{support}(f \circ \varphi) = \varphi^{-1}(\operatorname{support}(f)).$$

Applying Proposition 2.38 a second time, we see that $f \circ \varphi$ has compact support.

As a result, we can rewrite the formula to be proved in Theorem 3.34 as

$$\int_{\mathbb{R}^n} f(y)\,dy = \int_{\mathbb{R}^n} f(\varphi(x))|\det \varphi'(x)|\,dx,$$

and the supports will take care of themselves in the proof.

The result of Theorem 3.34 has to be regarded as preliminary. To understand the sense in which the result is limited, consider the case of polar coordinates in $\mathbb{R}^2$. In this case we can take

$$U = \left\{ \begin{pmatrix} r \\ \theta \end{pmatrix} \,\middle|\, 0 < r < +\infty \quad \text{and} \quad 0 < \theta < 2\pi \right\},$$

$$\varphi\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix},$$

and we have

$$\varphi(U) = \mathbb{R}^2 - \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \,\middle|\, x \geq 0 \right\}.$$

We readily compute that $\det \varphi'\binom{r}{\theta} = r$, and the desired formula is

$$\int_{\mathbb{R}^2} f(x, y)\,dx\,dy = \int_{0 \leq r < \infty,\ 0 \leq \theta < 2\pi} f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta.$$

At first glance this formula seems fine. But if we refer to the precise hypotheses, we see that f is assumed to vanish in a neighborhood of the set of points $(x, 0)$ with $x \geq 0$, as well as when (x, y) is sufficiently far from the origin. Without some sort of passage to the limit, the theorem therefore settles few cases of interest. This passage to the limit will be accomplished easily with the Lebesgue integral, and we therefore postpone the final form of the change-of-variables formula to Chapter VI.

In any event, we shall use Theorem 3.34 in proving the final change-of-variables formula, and thus a proof is warranted now. Before coming to the formal proof, it is well to understand the mechanism of the theorem. The proof will then flow easily from the analysis that is done for motivation.

The motivation for the theorem comes from taking f to be the constant function 1 and from thinking of φ as of the form $\varphi(y) = y_0 + L(y - y_0)$ with L linear. In $\mathbb{R}^3$, if we take U to be the cube $\{y = (y_1, y_2, y_3) \mid 0 \le y_i \le 1 \text{ for all } i\}$, along with $f = 1$, the formula asserts that $\varphi(U)$ has volume $|\det L|$. This is just the well-known fact about 3-by-3 matrices that the volume of the parallelepiped with sides u, v, w is the scalar $|(u \times v) \cdot w|$. For a corresponding result in $\mathbb{R}^n$, where vector product is not available, the relationship between the determinant and a volume has to be argued differently. One way of proceeding in $\mathbb{R}^n$ is to use row or column reduction to write the given matrix as the product of **elementary matrices** (those corresponding to the effect of a single step in the reduction), to check the change of variables for each factor, and to use the multiplication formula $\det(AB) = \det A \det B$ to obtain the result. This argument can be adjusted so as to work with a function f in place; the elementary matrices that interchange two variables are handled by Fubini's Theorem (Theorem 3.32 or Corollary 3.33), and the other elementary matrices are handled by the one-variable change-of-variables formula (Theorem 1.34).

That being the case, one can envision a proof of Theorem 3.34 that proceeds by approximation, using Taylor's Theorem (Theorem 3.11), at least if f is of class C^2. The contribution to the integrand from the integral remainder term in the Taylor expansion of φ is to be estimated as an error term. The approximation generates an additional error term because the image of U under φ does not match the image of U under the approximating first-order expansion of φ. Of course, one cannot expect the approximation to be very good far away from the point where the Taylor expansion is centered, and thus the argument needs to be carried out locally. The local contributions can then be pieced together by using a partition of unity. Such an argument can actually be carried out, but the argument is lengthy.

A more economical argument comes by finding a nonlinear analog of row or column reduction. The Inverse Function Theorem will allow us to prove that a general φ decomposes into suitably defined nonlinear elementary transformations, but the decomposition is valid only locally. A partition of unity is used to piece together the local results and obtain the theorem. We introduce two kinds of nonlinear elementary transformations:

(i) a **flip** β, which interchanges two coordinates. This is a linear function, and it satisfies $|\det \beta'(x)| = 1$ for all x. Application of Fubini's Theorem in the form of Corollary 3.33 shows that Theorem 3.34 is valid when φ is a flip.

(ii) a **primitive mapping**

$$\psi(x_1, \dots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ g(x_1, \dots, x_n) \\ x_{i+1} \\ \vdots \\ x_n \end{pmatrix},$$

where g is real-valued and occurs in a single entry. If that entry is the i^{th} entry, then the Jacobian matrix of ψ is the identity matrix except in the i^{th} row, where the entries are $\frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_n}$. Hence $|\det \psi'(x)| = \left|\frac{\partial g}{\partial x_i}\right|$. To prove Theorem 3.34 for a primitive mapping of this kind, it is enough to handle $i = 1$. If we write $x = (x_1, x')$ and $y = (y_1, x')$ with x' in $\mathbb{R}^{n-1}$, Fubini's Theorem (Corollary 3.33) reduces matters to showing that

$$\int_{\mathbb{R}^{n-1}} \Big[\int_{\mathbb{R}} f(y_1, x')\, dy_1 \Big]\, dx' = \int_{\mathbb{R}^{n-1}} \Big[\int_{\mathbb{R}} f(g(x_1, x'), x') \Big| \frac{\partial g}{\partial x_1}(x_1, x') \Big|\, dx_1 \Big]\, dx'$$

under suitable hypotheses on g, and it is enough to prove that the inner integrals are equal for all x'. Theorem 1.34 yields the equality of the inner integrals if g is a C^1 function for which $g(x_1, x')$ is defined for x_1 in an interval for any relevant x', and if $\left|\frac{\partial g}{\partial x_1}(x_1, x')\right|$ is everywhere positive at the points in question.

In the linear case a primitive mapping ψ for which $g(x)$ appears in the i^{th} entry is given by a matrix that is the identity except in the i^{th} row. For ψ' to be nonvanishing, the diagonal entry in the i^{th} row must be nonzero. This kind of matrix is not always elementary but is the product of n elementary matrices.

What needs to be proved for Theorem 3.34 is that apart from translations, any nonlinear φ as in Theorem 3.34 can be decomposed into the product of primitive transformations and flips, at least locally. The argument will peel primitive mappings from the right side of φ and flips from the left side. In that sense it will be a nonlinear version of column reduction with primitive mappings and row reduction with flips. The decomposition will be forced to be local because it uses the Inverse Function Theorem, which guarantees the existence of an inverse function only locally.

Lemma 3.35. Suppose that E is an open neighborhood of 0 in $\mathbb{R}^n$ and that $\varphi : E \to \mathbb{R}^n$ is a C^1 function such that $\varphi(0) = 0$ and $\varphi'(0)^{-1}$ exists. Then there is a subneighborhood of 0 in $\mathbb{R}^n$ in which φ factors as

$$\varphi = \beta_1 \circ \cdots \circ \beta_{n-1} \circ \psi_n \circ \cdots \circ \psi_1,$$

where each β_j is a flip or the identity and each ψ_j is a primitive C^1 function in some open neighborhood of 0 such that $\psi_j(0) = 0$ and $\psi'(0)^{-1}$ exists.

PROOF. Let us set up an inductive procedure by assuming at the start that

$$\varphi(x_1, \dots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_{i_0-1} \\ \varphi_{i_0}(x_1, \dots, x_n) \\ \vdots \\ \varphi_n(x_1, \dots, x_n) \end{pmatrix} \qquad (*)$$

with $1 \le i_0 \le n$. We shall make use of the following formula for multiplying two matrices A and B when B has the property that it is equal to the identity matrix except possibly in row i_0. The formula is

$$(AB)_{ij} = \begin{cases} A_{ii_0} B_{i_0 j} + A_{ij} B_{jj} & \text{if } j \ne i_0, \\ A_{ii_0} B_{i_0 i_0} & \text{if } j = i_0. \end{cases} \qquad (**)$$

It will be convenient to identify linear functions like $\varphi'(x)$ with their matrices, so that the $(i, j)^{\text{th}}$ entry $\varphi'(x)_{ij}$ of $\varphi'(x)$ is meaningful.

Let $j = j_0$ be the least row index for which the $(j, i_0)^{\text{th}}$ entry of $\varphi'(0)$ is nonzero. The index j_0 exists because $\varphi'(0)$ is nonsingular, and j_0 is $\ge i_0$ since the top $i_0 - 1$ rows of $\varphi'(x)$ match the corresponding rows of the identity matrix. Let

$$\beta_{i_0} = \begin{cases} \text{identity function} & \text{if } j_0 = i_0, \\ \text{flip of entries } j_0 \text{ and } i_0 & \text{if } j_0 > i_0. \end{cases}$$

Then $\beta_{i_0} \circ \varphi$ has the general form of $(*)$ except that the i_0^{th} and j_0^{th} entries have been interchanged. By inspection the Jacobian matrix at 0 of $\beta_{i_0} \circ \varphi$ equals the identity matrix in rows 1 through $i_0 - 1$ and has $(i_0, i_0)^{\text{th}}$ entry nonzero.

Thus if we possibly incorporate a composition with a flip into the definition of φ, we may assume that $\varphi'(0)_{i_0 i_0}$ is nonzero. Put

$$\psi(x_1, \dots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_{i_0-1} \\ \varphi_{i_0}(x_1, \dots, x_n) \\ x_{i_0+1} \\ \vdots \\ x_n \end{pmatrix}.$$

Then $\psi'(x)$ is an n-by-n matrix with

$$\psi'(x)_{ij} = \begin{cases} \delta_{ij} & \text{if } i \neq i_0, \\ \varphi'(x)_{i_0 j} & \text{if } i = i_0, \end{cases}$$

where δ_{ij} is the Kronecker delta. Since $\det \psi'(0) = \varphi'(0)_{i_0 i_0} \neq 0$, we can apply the Inverse Function Theorem (Theorem 3.17) to ψ, obtaining a C^1 inverse function ψ^{-1} that carries an open neighborhood of 0 onto an open subset of the domain of φ, has $\psi^{-1}(0) = 0$, and has derivative $(\psi^{-1})'(y) = \psi'(x)^{-1}$, where x and y are related by $y = \psi(x)$ and $x = \psi^{-1}(y)$. Using $(**)$, we readily verify that

$$(\psi'(x)^{-1})_{ij} = \begin{cases} \delta_{ij} & \text{if } i \neq i_0, \\ -(\varphi'(x)_{i_0 i_0})^{-1} \varphi'(x)_{i_0 j} & \text{if } i = i_0 \neq j, \\ (\varphi'(x)_{i_0 i_0})^{-1} & \text{if } i = j = i_0. \end{cases}$$

Therefore

$$((\psi^{-1})'(y))_{ij} = \begin{cases} \delta_{ij} & \text{if } i \neq i_0, \\ -(\varphi'(x)_{i_0 i_0})^{-1} \varphi'(x)_{i_0 j} & \text{if } i = i_0 \neq j, \\ (\varphi'(x)_{i_0 i_0})^{-1} & \text{if } i = j = i_0. \end{cases}$$

Form $\eta = \varphi \circ \psi^{-1}$. By the chain rule (Theorem 3.10), we have $\eta'(0) = \varphi'(0)(\psi^{-1})'(0)$, and this is nonsingular. Combining the formula for $((\psi^{-1})'(y))_{ij}$ with the chain rule and $(**)$ gives

$$\begin{aligned}
\eta'(x)_{ij} &= (\varphi'(x)(\psi^{-1})'(y))_{ij} \\
&= \begin{cases} \varphi'(x)_{i i_0} ((\psi^{-1})'(y))_{i_0 j} + \varphi'(x)_{ij} ((\psi^{-1})'(y))_{jj} & \text{if } j \neq i_0, \\ \varphi'(x)_{i i_0} ((\psi^{-1})'(y))_{i_0 i_0} & \text{if } j = i_0, \end{cases} \\
&= \begin{cases} \varphi'(x)_{i i_0} (-(\varphi'(x)_{i_0 i_0})^{-1} \varphi'(x)_{i_0 j}) + \varphi'(x)_{ij} & \text{if } j \neq i_0, \\ \varphi'(x)_{i i_0} (\varphi'(x)_{i_0 i_0})^{-1} & \text{if } j = i_0. \end{cases}
\end{aligned}$$

Since $\varphi'(x)_{i i_0}$ is 0 for $i < i_0$, the above formula shows that $\eta'(x)_{ij} = \delta_{ij}$ for $i < i_0$. For $i = i_0$, the formula shows first that $\eta'(x)_{i_0 j}$ is 0 for $j \neq i_0$ and then that $\eta'(x)_{i_0 j}$ is 1 for $j = i_0$. Thus $\eta'(x)_{ij} = \delta_{ij}$ for $i \leq i_0$. Consequently the i^{th} entry of $\eta(x)$ is $x_i + c_i$ if $i \leq i_0$, where c_i is a constant. Evaluating η at $x = 0$, we see that $c_i = 0$. Thus $\eta(x)$ has the same general shape as $(*)$ except that the i_0^{th} entry is now x_{i_0}.

Following this argument inductively for $i = 1, \ldots, n-1$ leads us to a decomposition

$$\eta = \beta_{n-1} \circ \cdots \circ \beta_1 \circ \varphi \circ \psi_1^{-1} \circ \cdots \circ \psi_{n-1}^{-1}, \tag{$\dagger$}$$

where each β_j is a flip or the identity and where each ψ_j is primitive. The function η has $\eta(0) = 0$ and $\eta'(0)$ nonsingular, and η has the form

$$\eta(x_1, \dots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ \xi(x_1, \dots, x_n) \end{pmatrix}.$$

Therefore η is primitive. Solving (†) for φ thus exhibits φ as decomposed into the required form.

PROOF OF THEOREM 3.34. We are to prove that

$$\int_{\varphi(U)} f(y)\,dy = \int_U f(\varphi(x))|\det \varphi'(x)|\,dx \tag{$*$}$$

whenever $\varphi : U \to \varphi(U)$ is a C^1 function between open sets with a C^1 inverse and $f : \varphi(U) \to \mathbb{R}$ is continuous and has compact support lying in $\varphi(U)$. In the argument we shall work with several functions in place of φ, and the set U may be different for each. We have seen that $(*)$ holds if φ is a flip or an invertible primitive function. Let us observe also that $(*)$ holds if φ is a translation $\varphi(x) = x + x_0$ for some x_0 in $\mathbb{R}^n$; the reason is that $(*)$ in this case can be reduced via successive uses of Fubini's Theorem (Corollary 3.33) to the 1-dimensional case, where we know it to be true by Theorem 1.34.

If $(*)$ holds when φ is either $\alpha : U \to \alpha(U)$ or $\beta : \alpha(U) \to \beta(\alpha(U))$, then $(*)$ holds when φ is the composition $\gamma = \beta \circ \alpha : U \to \beta(\alpha(U))$ because

$$\begin{aligned}
\int_{\mathbb{R}^n} f(z)\,dz &= \int_{\mathbb{R}^n} f(\beta(y))|\det \beta'(y)|\,dy \\
&= \int_{\mathbb{R}^n} f(\beta(\alpha(x))|\det \beta'(\alpha(x))||\det \alpha'(x)|\,dx \\
&= \int_{\mathbb{R}^n} f(\gamma(x))|\det(\beta'(\alpha(x))\alpha'(x))|\,dx \\
&= \int_{\mathbb{R}^n} f(\gamma(x))|\det \gamma'(x)|\,dx,
\end{aligned}$$

the last two steps holding by the formula $\det(BA) = \det B \det A$ and the chain rule (Theorem 3.10).

For any a in the given set U, Lemma 3.35 applies to the function φ_a carrying $U - a$ to $\varphi(U) - \varphi(a)$ and defined by $\varphi_a(x) = \varphi(x + a) - \varphi(a)$ because $\varphi_a(0) = 0$ and $\varphi_a'(0) = \varphi'(a)$. The lemma produces an open neighborhood E_a of

0 on which φ_a factors as a composition of flips and invertible primitive functions. If τ_{x_0} denotes the translation $\tau_{x_0}(x) = x + x_0$, then $\varphi_a = \tau_{-\varphi(a)} \circ \varphi \circ \tau_a$ shows that $\varphi = \tau_{\varphi(a)} \circ \varphi_a \circ \tau_{-a}$. Therefore φ factors on the open neighborhood $E_a + a$ as the composition of translations, flips, and invertible primitive functions. From the previous paragraph we conclude for each $a \in U$ that $(*)$ holds for φ if f is continuous and is compactly supported in the open neighborhood $\varphi(E_a + a)$ of $\varphi(a)$.

As a varies through U, the subsets $V_a = \varphi(E_a + a)$ of $\varphi(U)$ form an open cover of $\varphi(U)$. Fix f continuous with compact support K in $\varphi(U)$. By compactness a finite subfamily of the family $\{V_a\}$ forms an open cover of K. Applying Proposition 3.14, we obtain a finite family $\Psi = \{\psi\}$ of continuous functions defined on $\varphi(U)$ and taking values in [0, 1] with the properties that

(i) each ψ is 0 outside of some compact set contained in some V_a,
(ii) $\sum_{\psi \in \Psi} \psi$ is identically 1 on K.

Then property (i) and the conclusion of the previous paragraph show that $(*)$ holds for ψf. From (ii), we have $\sum_\psi \psi f = f$ on $\varphi(U)$. Since there are only finitely many terms in the sum, we can interchange sum and integral and conclude that $(*)$ holds for f. This completes the proof.

One final remark is appropriate: Theorem 3.34 immediately extends from the scalar-valued case as stated to the case that f takes values in $\mathbb{R}^m$ or $\mathbb{C}^m$.

11. Problems

1. Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. Prove that the Hilbert–Schmidt norm satisfies
   (a) $|TS| \le |T|\,|S|$ if S is in $L(\mathbb{F}^n, \mathbb{F}^m)$ and T is in $L(\mathbb{F}^m, \mathbb{F}^k)$,
   (b) $|1| = \sqrt{n}$ if $n = m$ and 1 denotes the identity function on $\mathbb{F}^n$.
2. Suppose that $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear function with Jacobian matrix A. What is $f'(x_0)$?
3. Suppose that $f : \mathbb{R}^2 \to \mathbb{R}^1$ has $|f(x)| \le |x|^2$ for all x. Prove that f is differentiable at $x = 0$.
4. Let $x = (x_1, \dots, x_n)$ and $u = (u_1, \dots, u_n)$ be in $\mathbb{R}^n$. For $f : \mathbb{R}^n \to \mathbb{R}$ differentiable at x, use the chain rule to derive a formula for $\frac{d}{dt} f(x + tu)\big|_{t=0}$.
5. Compute $\exp tX$ from the definition for $X = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$, and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

6. It was observed in Section 6 in the context of polar coordinates that the Implicit Function Theorem implies the Inverse Function Theorem. Namely, the pair of polar-coordinate formulas $(u, v) = (r\cos\theta, r\sin\theta)$ was inverted by applying the Implicit Function Theorem to the system of equations

$$r\cos\theta - u = 0, \qquad r\sin\theta - v = 0.$$

Using this example as a model, derive the Inverse Function Theorem in the general case from the Implicit Function Theorem in the general case.

7. Define $\int_1^\infty$ to mean $\lim_{N\to\infty}\int_1^N$ when the integrand is continuous. Prove or disprove:

$$\int_0^1 \Big[\int_1^\infty (e^{-xy} - 2e^{-2xy})\,dx\Big]\,dy = \int_1^\infty \Big[\int_0^1 (e^{-xy} - 2e^{-2xy})\,dy\Big]\,dx.$$

Problems 8–9 use Fubini's Theorem to supplement the theory of Fourier series as given in Section I.10.

8. Let f and g be continuous complex-valued periodic functions of period 2π, and define their **convolution** to be the function

$$f * g(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)g(t)\,dt.$$

(a) Show that $f * g$ is continuous periodic and that $f * g = g * f$.
(b) Let $f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$ and $g(x) \sim \sum_{n=-\infty}^{\infty} d_n e^{inx}$. Prove that $(f * g)(x) \sim \sum_{n=-\infty}^{\infty} c_n d_n e^{inx}$.
(c) Prove that the Fourier series of $f * g$ converges uniformly.

9. Let f, g, and h be continuous complex-valued periodic functions of period 2π. Prove that $f * (g * h) = (f * g) * h$.

Problems 10–13 deal with homogeneous functions. If $f : \mathbb{R}^n - \{0\} \to \mathbb{R}$ is a function not identically 0 such that $f(rx) = r^d f(x)$ for all x in $\mathbb{R}^n - \{0\}$ and all $r > 0$, we say that f is **homogeneous** of degree d. For example, the function in the first problem below is homogeneous of degree 0.

10. On $\mathbb{R}^2$, define

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

Prove that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist everywhere in $\mathbb{R}^2$ and that f is not continuous at $(0, 0)$.

11. Let $f : \mathbb{R}^n - \{0\} \to \mathbb{R}$ be smooth and homogeneous of degree d.
   (a) Prove that if $d = 0$, then $f(x)$ is bounded on $\mathbb{R}^n - \{0\}$ and that f extends to be continuous at 0 only if it is constant.
   (b) Prove that if $d > 0$, then the definition $f(0) = 0$ makes f continuous for all x in $\mathbb{R}^n$, while if $d < 0$, then no definition of $f(0)$ makes f continuous at 0.
   (c) Prove that $\frac{\partial f}{\partial x_j}$ is homogeneous of degree $d - 1$ unless it is identically 0.
   (d) If f is homogeneous of degree 1 and satisfies $f(-x) = -f(x)$ and $f(0) = 0$, prove that each $\frac{\partial f}{\partial x_j}$ exists at 0 but that $\frac{\partial f}{\partial x_j}$ is not continuous at 0 unless it is constant.

12. On $\mathbb{R}^2$, let f be the function homogeneous of degree 1 given by

$$f(x, y) = \begin{cases} \dfrac{x^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

   (a) Prove that f is continuous at $(0, 0)$.
   (b) Prove that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist at $(0, 0)$ but are not continuous there.
   (c) Calculate $\frac{d}{dt} f(t + tu)\big|_{t=0}$ for $x = 0$ and $u = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$. Show that the formula in Problem 4 fails, and conclude that f is not differentiable at $(0, 0)$.

13. On $\mathbb{R}^2$, let f be the function homogeneous of degree 2 given by

$$f(x, y) = \begin{cases} \dfrac{xy(x^2 - y^2)}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

   (a) Prove that f, $\frac{\partial f}{\partial x}$, and $\frac{\partial f}{\partial y}$ are continuous on all of $\mathbb{R}^2$.
   (b) Prove that $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ exist at $(0, 0)$ but are not continuous there.
   (c) Prove that $\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1$ and $\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1$.

Problems 14–15 concern "harmonic functions" in $\{(x, y) \in \mathbb{R}^2 \mid |(x, y)| < 1\}$, the **open unit disk** of the plane. A **harmonic function** is a complex-valued C^2 function satisfying the **Laplace equation** $\Delta u(x, y) = 0$, where Δ is the **Laplacian** $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$.

14. If (r, θ) are regarded as polar coordinates, prove for all integers n that each function $r^{|n|} e^{in\theta}$ is a C^∞ function in the open unit disk and is harmonic there. Deduce that if $\{c_n\}$ is a doubly infinite sequence such that $\sum_{n=-\infty}^{\infty} c_n r^{|n|} e^{in\theta}$ converges absolutely for each r with $0 \le r < 1$, then the sum is a C^∞ function in the open unit disk and is harmonic there.

15. Prove that if u is harmonic in the unit disk, then so is the function $u \circ R$, where R is the rotation about the origin given by $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

Problems 16–20 illustrate the Inverse and Implicit Function Theorems.

16. Verify that the equations $u = x^4y + x$ and $v = x + y^3$ define a function from $\mathbb{R}^2$ to $\mathbb{R}^2$ whose derivative at $(1, 1)$ is given by the matrix $\begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}$. This matrix being invertible, the Inverse Function Theorem applies. Let the locally defined C^1 inverse function be given by $x = F(u, v)$ and $y = G(u, v)$ in an open neighborhood of $(u, v) = (2, 2)$, the point $(2, 2)$ having the property that $F(2, 2) = 1$ and $G(2, 2) = 1$. Find $\frac{\partial F}{\partial u}(2, 2)$.

17. Show that the equations
$$\begin{aligned} x^2 - y\cos(uv) + z^2 &= 0, \\ x^2 + y^2 - \sin(uv) + 2z^2 &= 2, \\ xy - \sin u \cos v + z &= 0, \end{aligned}$$
implicitly define x, y, z as C^1 functions of (u, v) near $x = 1$, $y = 1$, $u = \pi/2$, $v = 0$, and $z = 0$, and find $\frac{\partial x}{\partial u}$ and $\frac{\partial x}{\partial v}$ for the function $x(u, v)$. Is the function $x(u, v)$ of class C^∞?

18. Regard the operation of squaring an n-by-n matrix as a function from $\mathbb{R}^{n^2}$ to $\mathbb{R}^{n^2}$, and show that this mapping is invertible on some open set of the domain that contains the identity matrix.

19. **(Lagrange multipliers)** Let f and g be real-valued C^1 functions defined on an open subset U of $\mathbb{R}^n$, and let $S = \{x \in U \mid g(x) = 0\}$. Prove that if $f\big|_S$ has a local maximum or minimum at a point x_0 of S, then either $g'(x_0) = 0$ or there exists a number λ such that $f'(x_0) + \lambda g'(x_0) = 0$.

20. **(Arithmetic-geometric mean inequality)** Using Lagrange multipliers, prove that any n real numbers $a_1, \dots, a_n$ that are ≥ 0 satisfy
$$\sqrt[n]{a_1 a_2 \cdots a_n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

CHAPTER IV

Theory of Ordinary Differential Equations and Systems

Abstract. This chapter treats the theory of ordinary differential equations, both linear and nonlinear.

Sections 1–4 establish existence and uniqueness theorems for ordinary differential equations. The first section gives some examples of first-order equations, mostly nonlinear, to illustrate certain kinds of behavior of solutions. The second section shows, in the presence of continuity for a vector-valued F satisfying a "Lipschitz condition," that the first-order system $y' = F(t, y)$ has a unique local solution satisfying an initial condition $y(t_0) = y_0$. Since higher-order equations can always be reduced to first-order systems, these results address existence and uniqueness for n^{th}-order equations as a special case. Section 3 shows that the solutions to a system depend well on the initial condition and on any parameters that are present in F. Section 4 applies these results to existence of integral curves for a vector field and to construction of coordinate systems from families of integral curves.

Sections 5–8 concern linear systems. Section 5 shows that local solutions of linear systems may be extended to global solutions and that in the homogeneous case the vector space of global solutions has dimension equal to the size of the system. The method of variation of parameters reduces the solution of any linear system to the solution of a homogeneous linear system. Sections 6–7 identify explicit solutions to n^{th}-order linear equations and first-order linear systems. The "Jordan canonical form" of a square matrix plays a role in the case of a system. Section 8 discusses power-series solutions to second-order homogeneous linear equations whose coefficients are given by convergent power series, as well as solutions that arise in the case of regular singular points. Two kinds of special functions are mentioned that result from this study—Legendre polynomials and Bessel functions.

1. Qualitative Features and Examples

To introduce the subject of ordinary differential equations, this section gives examples of some qualitative features and complicated phenomena that can occur in such equations.

If F is a complex-valued function of $n+2$ variables, a function $y(t)$ is said to be a **solution** of the **ordinary differential equation**

$$F(t, y, y', y'', \dots, y^{(m)}) = 0$$

of m^{th} **order** on the open interval (a, b) if

$$F(t, y(t), y'(t), \dots, y^{(m)}(t)) = 0$$

identically for $a < t < b$. The equation is "ordinary" in the sense that there is only one independent variable. The equation is said to be **linear** if it is of the form

$$a_m(t)y^{(m)} + a_{m-1}(t)y^{(m-1)} + \cdots + a_1(t)y' + a_0(t)y = q(t),$$

and it is **homogeneous linear** if in addition, q is the 0 function. A linear ordinary differential equation has **constant coefficients** if $a_m(t), \ldots, a_0(t)$ are all constant functions.

Let us come to examples, which will point toward the enormous variety of phenomena that can occur. We stick to the first-order case, and all the examples will have F real-valued. Let us look only for real-valued solutions. Pictures indicating the qualitative behavior of the solutions of each of the examples are in Figure 4.1.

EXAMPLES.

(1) Simple equations can have relatively complicated solutions. This is already true for the equation

$$y' = 1/t \qquad \text{on the interval } (0, +\infty).$$

Integration shows that all solutions are of the form $\log t + c$; on an interval of negative t's, the solutions are of the form $\log|t| + c$. The c comes from a corollary of the Mean Value Theorem that says that a real-valued function on an open interval with 0 derivative everywhere is necessarily constant.[1] Another example, but with no singularity, is $y' = ty$. To solve this equation on intervals where $y(t) \neq 0$, write $y'/y = t$, so that $\log|y| = \frac{1}{2}t^2 + a$ and $|y| = e^a e^{t^2/2}$. Thus $y(t) = ce^{t^2/2}$, with $c \neq 0$ constant, on any interval where $y(t)$ is nowhere 0. The function $y(t) = 0$ is a solution as well, and all real solutions on an interval are of the form $y(t) = ce^{t^2/2}$ with c real. See Figures 4.1a and 4.1b.

(2) Solutions may not be defined on obvious intervals. For the equation

$$ty' + y = \sin t,$$

we can recognize the two sides as $\frac{d}{dt}(ty)$ and $\frac{d}{dt}(-\cos t)$. Therefore $ty = c - \cos t$. Dividing by t, we obtain $y(t) = \frac{c-\cos t}{t}$ on any interval that does not contain 0. What about intervals containing $t = 0$? If we put $t = 0$ in the formula $ty = c - \cos t$, we see that c must be 1. In this case we can define $y(0) = 0$ there, and then $y'(0)$ exists. We obtain the additional solution

$$y(t) = \begin{cases} \dfrac{1 - \cos t}{t} & \text{for } t \neq 0, \\ 0 & \text{for } t = 0, \end{cases}$$

on any open interval containing 0. Figure 4.1c shows graphs of some solutions.

[1] See Section A2 of the appendix for further information.

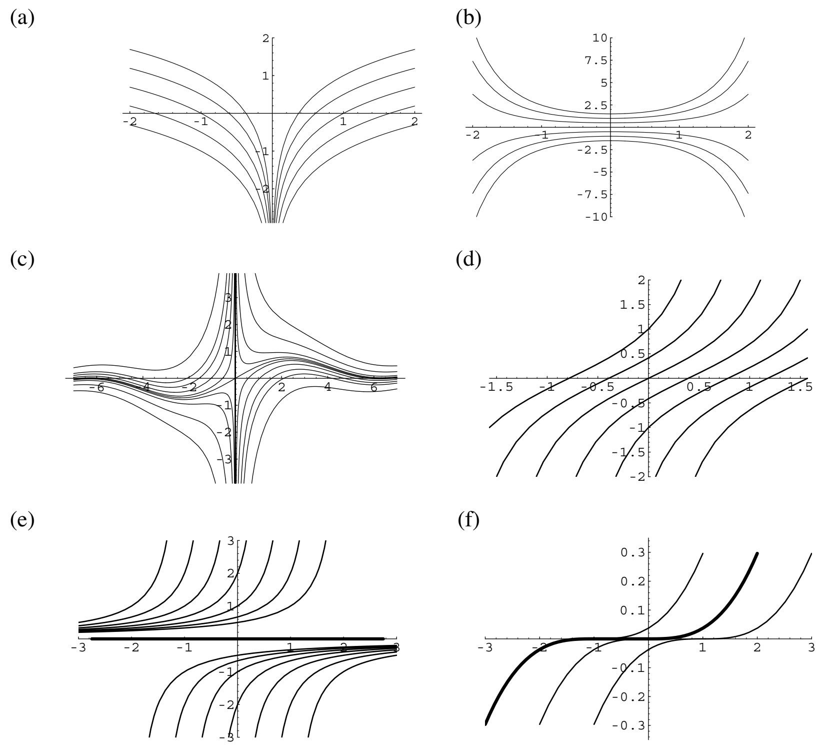


FIGURE 4.1. Graphs of solutions of some first-order ordinary differential equations: (a) $y' = 1/t$, (b) $y' = ty$, (c) $ty' + y = \sin t$, (d) $y' = y^2 + 1$, (e) $y' = y^2$, (f) $y' = y^{2/3}$.

(3) Even if the equation seems nice for all t, the solutions may not exist for all t. An example occurs with

$$y' = y^2 + 1,$$

which we solve by the steps $\frac{d}{dt}(\arctan y) = 1$, $\arctan y = t + c$, $y = \tan(t + c)$. The solutions behave badly when $t + c$ is any odd multiple of $\pi/2$. Solutions are defined at most on intervals of length π. Figure 4.1d shows graphs of some solutions for this example.

(4) Some solutions may look quite different from all the others. For example, with

$$y' = y^2,$$

we solve by $-1/y = t + c$ for $y \neq 0$, so that $y(t) = -\dfrac{1}{t + c}$. Also, $y(t) = 0$ is a solution. Here the solutions of the form $y(t) = -\frac{1}{t+c}$ are not defined for all t, but the solution $y(t) = 0$ *is* defined for all t. We might think of $y(t) = 0$ as the limiting case with c tending to $\pm\infty$. Figure 4.1e shows graphs of some of the solutions for this example.

(5) New solutions can sometimes be pieced together from old ones. For example, the equation

$$y' = y^{2/3}$$

is solved where $y \neq 0$ by the steps $y^{-2/3}y' = 1$, $3y^{1/3} = t + c$, and $y(t) = \frac{1}{27}(t + c)^3$. But also $y(t) = 0$ is a solution. In fact, we can piece solutions of these types together. For example, the function

$$y(t) = \begin{cases} \frac{1}{27}(t+1)^3 & \text{for } t < -1, \\ 0 & \text{for } -1 \leq t \leq 0, \\ \frac{1}{27}t^3 & \text{for } 0 < t, \end{cases}$$

is a solution on $(-\infty, +\infty)$. Figure 4.1f shows graphs of some of the solutions for this example.

One thing that stands out in the above examples is that the set of solutions seems to depend, more or less, on a single parameter c. The inference is that nothing much worse than the c occurs because somewhere an integration is taking place and the Mean value Theorem is controlling how many indefinite integrals there can be. One way of trying to quantify this statement about how the number of solutions is limited is to say that for any fixed $t = t_0$ and given real number y_0, there is only one solution $y(t)$ near t_0 with $y(t_0) = y_0$. This statement is not quite accurate, however, as Example 5 shows. The uniqueness theorem in Section 2 will give a precise result. The data (t_0, y_0) are called an **initial condition**.

Something else that stands out, although perhaps not without the visual aid of the graphs of solutions as in Figure 4.1, is that the graphed solutions appear to fill the entire part of the plane corresponding to the t's under study. In the framework of the previous paragraph, the statement is that for any fixed $t = t_0$ and given real number y_0, there exists a solution $y(t)$ near t_0 with $y(t) = y_0$. The existence theorem in Section 2 will give a precise result.

WEAK VERSION OF EXISTENCE AND UNIQUENESS THEOREMS. Let D be a nonempty convex open set in $\mathbb{R}^2$, and let (t_0, y_0) be in E. If $F : D \to \mathbb{R}$ is a continuous function such that $\frac{\partial}{\partial y}F(t, y)$ exists and is continuous in D, then for any sufficiently small open interval of t's containing t_0, the equation $y' = F(t, y)$ has a unique solution $y(t)$ with $y(t_0) = y_0$ such that the graph of $t \mapsto y(t)$ lies in D.

An improved theorem, together with a proof, will be given in Section 2. The proof of existence uses "Picard iterations," and the idea is as follows. First we convert the differential equation into an equivalent integral equation

$$y(t) = \int_{t_0}^{t} F(s, y(s))\, ds + y_0.$$

Second we use the right side as input and the left side as output to define successive approximations to a solution:

$$\begin{aligned} y_0(t) &= y_0, \\ y_1(t) &= \int_{t_0}^{t} F(s, y_0(s))\, ds + y_0, \\ &\vdots \\ y_{n+1}(t) &= \int_{t_0}^{t} F(s, y_n(s))\, ds + y_0. \end{aligned}$$

Third we use the Weierstrass M test to show that the series with partial sums $y_N(t) = y_0 + \sum_{n=1}^{N} (y_n(t) - y_{n-1}(t))$ is uniformly convergent. If the limiting function is denoted by $y(t)$, we check that $y(t)$ satisfies the integral equation from which we started. Hence $y(t)$ is a solution of the differential equation.

2. Existence and Uniqueness

In this section we state and prove the main existence and uniqueness theorems for solutions of ordinary differential equations. First let us establish an appropriate setting more general than the one in Section 1.

The examples in Section 1 were all of the first order. They could all have been written in the form $y = F(t, y)$ with F real-valued, and we considered real-valued solutions $y(t)$. From equations as simple as $y'' + y' + y = 0$, whose real-valued solutions are

$$y(t) = a_1 e^{-t/2} \cos(t\sqrt{3}/2) + a_2 e^{-t/2} \sin(t\sqrt{3}/2),$$

we know that it can be easier to work, at least initially, with complex-valued solutions. In this particular case, it is easier as a first step to find all complex-valued solutions, namely

$$y(t) = c_1 \exp\left(\tfrac{1}{2}(-1 + i\sqrt{3})t\right) + c_2 \exp\left(\tfrac{1}{2}(-1 - i\sqrt{3})t\right),$$

and then to extract the real-valued solutions from them. The solution method, which will be discussed in more detail in Section 6 below, involves finding all complex solutions of a certain polynomial equation with real coefficients, and the method is more natural if the coefficients of the polynomial equation are allowed to be complex.

Thus right away, it is natural to consider first-order equations $y' = F(t, y)$ with F complex-valued and to look for complex-valued solutions. The theory in Chapter III avoided working with functions of several variables in which some of the variables are complex, and we can update the theory of Chapter III here. The technique, which is to consider the complex variable y as two real variables $\operatorname{Re} y$ and $\operatorname{Im} y$, is again applicable. Thus we have only to think of $F(t, y)$ as a function of three real variables, even if we do not separate y into its two components in writing $F(t, y)$, and the theory of Chapter III applies directly. In adopting the point of view that y is actually two real variables, we need to apply the same consideration to y', and we are led to view $y' = F(t, y)$ as a *system* of two simultaneous equations, namely $\operatorname{Re} y' = \operatorname{Re} F(t, y)$ and $\operatorname{Im} y' = \operatorname{Im} F(t, y)$. This viewpoint merely makes our functions conform to the prescriptions of Chapter III. It is not necessary to work with the expanded notation; all we have to remember is that in this part of the theory we never differentiate a function with respect to a *complex* variable.

The utility of allowing $y' = F(t, y)$ to represent a *system* of ordinary differential equations has, in any event, been thrust upon us. Let us consider the notion of a system a bit more. With a little trick the second-order equation $y'' + y' + y = 0$ can itself be transformed into a system, quite apart from the issue of real vs. complex variables. The trick is to introduce two unknown functions u_1 and u_2 to play the roles of y and y'. Then u_1 and u_2 satisfy $u_2 = u_1'$ and $u_2' = u_1'' = y'' = -y' - y = -u_2 - u_1$. In other words, u_1 and u_2 satisfy the system

$$\begin{aligned} u_1' &= u_2, \\ u_2' &= -u_1 - u_2. \end{aligned}$$

Conversely if $u_1(t)$ and $u_2(t)$ satisfy this system of equations, then $y(t) = u_1(t)$ is a solution of $y'' + y' + y = 0$. In this way, the given second-order equation is completely equivalent to a certain system of two first-order equations with two unknown functions.

Let F be a function defined on an open set D of $\mathbb{R} \times \mathbb{C}^{km}$ and taking values in $\mathbb{C}^k$. A $\mathbb{C}^k$-valued function $y(t) = (y_1(t), \dots, y_k(t))$ is said to be a **solution** of the **system** $F(t, y, y', \dots, y^{(m)}) = 0$ of k ordinary differential equations of **order** m in the open interval (a, b) if $F(t, y(t), y'(t), \dots, y^{(m)}) = 0$ identically for $a < t < b$.

We saw that the single second-order equation $y'' + y' + y = 0$ is equivalent to a certain first-order system of two equations, and the technique for exhibiting this equivalence works more generally: a system of k equations of order m that has been solved for the m^{th}-order derivatives is equivalent to a system of km equations of first order.

We shall consider first-order systems of the form $y' = F(t, y)$, where F is continuous on an open subset D of $\mathbb{R} \times \mathbb{C}^n$ and takes values in $\mathbb{C}^n$. The example $y' = y^{2/3}$ in Section 1 fits these hypotheses, and we saw that the hoped-for uniqueness fails for this equation. In the weak theorem stated at the end of Section 1, an additional hypothesis was imposed in order to address this problem: for $y' = F(t, y)$ with only real-valued solutions of interest, the hypothesis is that $\partial F/\partial y$ exists and is continuous on the domain D of F. Generalizing this condition presumably means saying something about partial derivatives in each of the directions y_j for $1 \leq j \leq n$. In addition, we must remember the injunction against differentiating with respect to complex variables. Thus we really expect a condition concerning $2n$ first-order derivatives. Fortunately there is an easily stated less-stringent condition that is nevertheless good enough. The condition is that F satisfy a **Lipschitz condition** in its y variable, i.e., that there exist a real number k such that

$$|F(t, y_1) - F(t, y_2)| \leq k|y_1 - y_2|$$

for all pairs of points (t, y_1) and (t, y_2) in the domain D of F.

If F is a real-valued continuous function of two real variables with a continuous partial derivative in the second variable, then the Mean Value Theorem gives

$$F(t, y_1) - F(t, y_2) = (y_1 - y_2)\,\frac{\partial F}{\partial y}(t, \xi)$$

with ξ between y_1 and y_2, *provided* the line segment from (t, y_1) to (t, y_2) lies in the domain D of F. The partial derivative is bounded on any compact subset of D, and thus F satisfies, on any compact convex subset of D, a Lipschitz condition in the second variable.

Theorem 4.1 (Picard–Lindelöf Existence Theorem). Let D be a nonempty open set in $\mathbb{R}^1 \times \mathbb{C}^n$, let (t_0, y_0) be in D, and suppose that $F : D \to \mathbb{C}^n$ is a continuous function such that $F(t, y)$ satisfies a Lipschitz condition in the y variable and has $|F(t, y)| \leq M$ on D. Let R be a compact set in $\mathbb{R}^1 \times \mathbb{C}^n$ of the form

$$R = \big\{(t, y) \,\big|\, |t - t_0| \leq a \text{ and } |y - y_0| \leq b\big\},$$

and suppose that R is contained in D. Put $a' = \min\{a,\ b/M\}$. Then there exists a solution $y(t)$ of the system

$$y' = F(t, y)$$

on the open interval $|t - t_0| < a'$ satisfying the initial condition

$$y(t_0) = y_0.$$

REMARKS. A variant of Theorem 4.1 takes D to be in $\mathbb{R}^1 \times \mathbb{C}^n$ but insists only on continuity of F, not on the Lipschitz condition. Then a local solution still exists for $|t-t_0| < a'$. This better result, known as the "Cauchy–Peano Existence Theorem," appears in Problems 20–25 at the end of the chapter and is proved by an argument using Ascoli's Theorem. However, Example 5 in Section 1 shows that there is no corresponding uniqueness theorem, and within the text we omit the proof of the better existence theorem. Another variant of Theorem 4.1 assumes that the domain D of a given $F_{\mathbb{R}}$ lies in $\mathbb{R}^1 \times \mathbb{R}^n$, $F_{\mathbb{R}}$ takes values in $\mathbb{R}^n$, and y_0 is in $\mathbb{R}^n$. Then $y' = F_{\mathbb{R}}(t, y)$ has a solution $y(t)$ such that $y(t_0) = y_0$ and the range of y is $\mathbb{R}^n$. In fact, when $F_{\mathbb{R}}$ satisfies a Lipschitz condition in the y variable, this variant is a consequence of Theorem 4.1 as stated. To derive this variant, one extends the given function $F_{\mathbb{R}}$ from the subset of $\mathbb{R}^1 \times \mathbb{R}^n$ to a subset of $\mathbb{R}^1 \times \mathbb{C}^n$ by making it constant in $\operatorname{Im} y$. Specifically the new system is $y' = F(t, y)$ with $F(t, y) = F_{\mathbb{R}}(t, \operatorname{Re} y)$, and the initial condition remains as $y(t_0) = y_0$. The part of the system corresponding to equations for $\operatorname{Im} y'$ is just $\operatorname{Im} y' = 0$, since F is real-valued, and therefore $\operatorname{Im} y(t)$ is constant. Since y_0 is real, $\operatorname{Im} y(t)$ must be 0. Thus Theorem 4.1 yields a solution $y(t)$ with range $\mathbb{R}^n$ under these special hypotheses.

PROOF. The first step is to see that the set of differentiable functions $t \mapsto y(t)$ on $|t - t_0| < a'$ satisfying $y' = F(t, y)$ and $y(t_0) = y_0$ is the same as the set of continuous functions $t \mapsto y(t)$ on $|t - t_0| < a'$ satisfying the integral equation $y(t) = \int_{t_0}^t F(s, y(s))\, ds + y_0$.

If y is differentiable and satisfies the differential equation and the initial condition, then y is certainly continuous and hence $s \mapsto F(s, y(s))$ is continuous. Then $\int_{t_0}^t F(s, y(s))\, ds$ is differentiable by the Fundamental Theorem of Calculus (Theorem 1.32), and the differential equation shows that $y(t)$ and $\int_{t_0}^t F(s, y(s))\, ds$ have the same derivative for $|t - t_0| < a'$. Thus they differ by a constant. The constant is checked by putting $t = t_0$, and indeed y satisfies the integral equation.

Conversely if y is continuous and satisfies the integral equation, then $s \mapsto F(s, y(s))$ is continuous, and the Fundamental Theorem of Calculus shows that $\int_{t_0}^t F(s, y(s))\, ds$ is differentiable. This function equals $y(t) - y_0$ by the integral equation, and hence y is differentiable. Differentiating the two sides of the integral equation, we see that y satisfies the differential equation. Also, if we put $t = t_0$ in the integral equation, we see that y satisfies the initial condition $y(t_0) = y_0$.

Thus it is enough to prove existence for a continuous solution of the integral equation. For $t_0 - a' \le t \le t_0 + a'$, define inductively

$$
\begin{aligned}
y_0(t) &= y_0, \\
y_1(t) &= y_0 + \int_{t_0}^t F(s, y_0(s))\, ds,
\end{aligned}
$$

$$\vdots$$

$$y_n(t) = y_0 + \int_{t_0}^{t} F(s, y_{n-1}(s))\, ds,$$

with the usual convention that $\int_{t_0}^{t} = -\int_{t}^{t_0}$. Let us see inductively that the graph of $y_n(t)$ lies in the set

$$R' = \{(t, y) \mid |t - t_0| \le a' \text{ and } |y - y_0| \le b\},$$

for $|t - t_0| \le a'$. The graph of $y_0(t) = y_0$ is just $\{(t, y_0) \mid |t - t_0| < a'\}$, and this lies in R'. The inductive hypothesis is that $(t, y_{n-1}(t))$ lies in R' for $\{(t, y_0) \mid |t - t_0| \le a'\}$. Then

$$|y_n(t) - y_0| = \Big| \int_{t_0}^{t} F(s, y_{n-1}(s)\, ds \Big| \le M|t - t_0| \le Ma' \le b,$$

and therefore $(t, y_n(t))$ lies in R' for $|t - t_0| \le a'$. This completes the induction, and hence the graph of $y_n(t)$ lies in R' for $|t - t_0| \le a'$.

Now write

$$y_N(t) = y_0(t) + \sum_{n=1}^{N} [y_n(t) - y_{n-1}(t)]$$

for $N \ge 0$. We shall use the Weierstrass M test (Proposition 1.20), adapted to a series of functions with values in $\mathbb{C}^n$, to prove uniform convergence of this series. Thus we are to bound $|y_n(t) - y_{n-1}(t)|$, and we shall do so inductively for $n \ge 1$. We start from the inequality $|F(t, y)| \le M$ on R' and the Lipschitz condition

$$|F(t, y_j(t) - F(t, y_{j-1})| \le k|y_j(t) - y_{j-1}(t)| \quad \text{for } j \ge 1.$$

Say that $t_0 \le x \le t_0 + a'$ for definiteness. Then

$$|y_1(t) - y_0(t)| = \Big| \int_{t_0}^{t} F(s, y_0(s))\, ds \Big| \le M(t - t_0)$$

and

$$\begin{aligned} |y_2(t) - y_1(t)| &= \Big| \int_{t_0}^{t} [F(s, y_1(s)) - F(s, y_0(s))]\, ds \Big| \\ &\le \int_{t_0}^{t} |F(s, y_1(s)) - F(s, y_0(s))|\, ds \end{aligned}$$

$$
\begin{aligned}
&\leq \int_{t_0}^{t} k|y_1(s) - y_0(s)|\, ds \\
&\leq \int_{t_0}^{t} kM(s - t_0)\, ds \quad \text{from the previous display} \\
&= \frac{Mk(t - t_0)^2}{2!}.
\end{aligned}
$$

Now we carry out an induction. The base case is the estimate carried out above for $|y_1(t)-y_0(t)|$. The estimate for $|y_2(t)-y_1(t)|$ suggests the inductive hypothesis, namely the inequality

$$|y_{n-1}(t) - y_{n-2}(t)| \leq \frac{Mk^{n-2}(t - t_0)^{n-1}}{(n-1)!}.$$

Then we have

$$
\begin{aligned}
|y_n(t) - y_{n-1}(t)| &\leq \int_{t_0}^{t} |F(s, y_{n-1}(s) - F(t, y_{n-2}(s))|\, ds \\
&\leq \int_{t_0}^{t} k|y_{n-1}(s) - y_{n-2}(s)|\, ds \\
&\leq Mk^{n-1} \int_{t_0}^{t} \frac{(s - t_0)^{n-1}}{(n-1)!}\, ds \qquad \text{by inductive hypothesis} \\
&= \frac{Mk^{n-1}(t - t_0)^n}{n!},
\end{aligned}
$$

and the induction is complete. The argument when $t_0 - a' \leq t \leq t_0$ is completely similar, and the form of the estimate for the two cases combined is

$$|y_n(t) - y_{n-1}(t)| \leq \frac{Mk^{n-1}|t - t_0|^n}{n!} \qquad \text{for } |t - t_0| \leq a'.$$

There is no harm in assuming that k is > 0, and consequently

$$|y_n(t) - y_{n-1}(t)| \leq \frac{M}{k} \frac{k^n(a')^n}{n!}$$

independently of t. Since $\sum_{n=0}^{\infty} (n!)^{-1} k^n (a')^n = e^{ka'}$ is finite, the M test applies and shows that our series converges uniformly.

Thus $y_N(t)$ converges uniformly for $|t - t_0| \leq a'$, necessarily to a continuous function. We call this function $y(t)$. For $|t - t_0| \leq a'$, we have

$$
\begin{aligned}
\int_{t_0}^{t} F(s, y(s))\, ds &= \int_{t_0}^{t} [F(s, y(s)) - F(t, y_N(s))]\, ds + \int_{t_0}^{t} F(s, y_N(s))\, ds \\
&= \int_{t_0}^{t} [F(s, y(s)) - F(s, y_N(s))]\, ds + y_{N+1}(t) - y_0.
\end{aligned}
$$

On the right side, we have $\lim_N[y_{N+1}(t) - y_0] = y(t) - y_0$. Because of the Lipschitz condition the absolute value of the first term on the right side is

$$\leq a'k \sup_{|t-t_0|\leq a'} |y(t) - y_N(t)|,$$

and this tends to 0 as n tends to infinity. Thus

$$\int_{t_0}^{t} F(s, y(s))\, ds = y(t) - y_0,$$

and $y(t)$ is a continuous solution of the integral equation.

Theorem 4.2 (uniqueness theorem). Let D be a nonempty open set in $\mathbb{R}^1 \times \mathbb{C}^n$, let (t_0, y_0) be in D, and suppose that $F : D \to \mathbb{C}^n$ is a continuous function such that $F(t, y)$ satisfies a Lipschitz condition in the y variable. For any $a'' > 0$, there exists at most one solution $y(t)$ to the system

$$y' = F(t, y)$$

on the open interval $|t - t_0| < a''$ satisfying the initial condition

$$y(t_0) = y_0.$$

PROOF. As in the proof of Theorem 4.1, it is enough to prove uniqueness for the integral equation. Suppose that $y(t)$ and $z(t)$ are two solutions for $|t - t_0| < a''$. Fix $\epsilon > 0$. Then $|y(t) - z(t)|$ is bounded by some constant C for $|t - t_0| \leq a'' - \epsilon$, and F is assumed to satisfy a Lipschitz condition $|F(t, y_1) - F(t, y_2)| \leq k|y_1 - y_2|$ on D.

We argue as in the proof of Theorem 4.1, working first for $t_0 \leq t$ and starting from

$$|y(t) - z(t)| \leq C$$

and from

$$\begin{aligned}
|y(t) - z(t)| &= \Big| \int_{t_0}^{t} [F(s, y(s)) - F(s, z(s))]\, ds \Big| \\
&\leq \int_{t_0}^{t} |F(s, y(s)) - F(s, z(s))|\, ds \\
&\leq \int_{t_0}^{t} k|y(s) - z(s)|\, ds \\
&\leq Ck(t - t_0).
\end{aligned}$$

Inductively we suppose that

$$|y(t) - z(t)| \le \frac{Ck^{n-1}(t-t_0)^{n-1}}{(n-1)!}.$$

Then

$$\begin{aligned} |y(t) - z(t)| &\le \int_{t_0}^{t} |F(s, y(s)) - F(s, z(s))|\,ds \\ &\le \int_{t_0}^{t} k|y(s) - z(s)|\,ds \\ &\le Ck^n \int_{t_0}^{t} \frac{(s-x_0)^{n-1}}{(n-1)!}\,ds = \frac{Ck^n(t-t_0)^n}{n!}, \end{aligned}$$

and thus $|y(t) - z(t)| \le C(n!)^{-1}k^n(t-t_0)^n$ for all n. A similar estimate is valid for $t \le t_0$, and the combined estimate is

$$|y(t) - z(t)| \le \frac{Ck^n|t-t_0|^n}{n!}.$$

Since $\sum C(n!)^{-1}k^n|t-t_0|^n$ converges, the individual terms tend to 0. Therefore $y(t) = z(t)$ for $|t-t_0| \le a'' - \epsilon$. Since ϵ is arbitrary, $y(t) = z(t)$ for $|t-t_0| < a''$.

3. Dependence on Initial Conditions and Parameters

In abstract settings where the existence and uniqueness theorems play a role, it is frequently of interest to know how the unique solution depends on the initial data (t_0, y_0) such that $y(t_0) = y_0$. To quantify this dependence, let us write the unique solution corresponding to $y' = F(t, y)$ as $y(t, t_0, y_0)$ rather than $y(t)$. We continue to use y' to indicate the derivative in the t variable even though the differentiation is now actually a partial derivative.

Theorem 4.3. Let D be a nonempty open set in $\mathbb{R}^1 \times \mathbb{C}^n$, let (t, y^*) be in D, and suppose that $F : D \to \mathbb{C}^n$ is a continuous function such that $F(t, y)$ satisfies a Lipschitz condition in the y variable. Let R be a compact set in $\mathbb{R}^1 \times \mathbb{C}^n$ of the form

$$R = \big\{(t, y) \;\big|\; |t - t^*| \le a \text{ and } |y - y^*| \le b\big\},$$

suppose that R is contained in D, and let M be an upper bound for $|F|$ on R. Put $a' = \min\{a, b/M\}$. If $|t_0 - t^*| < a'/2$ and $|y_0 - y^*| < b/2$, then there exists a unique solution $t \mapsto y(t, t_0, y_0)$ on the interval $|t - t_0| < a'/2$ to the system and initial data

$$y' = F(t, y) \quad \text{and} \quad y(t_0, t_0, y_0) = y_0,$$

and the function $(t, t_0, y_0) \mapsto y(t, t_0, y_0)$ is continuous on the open set

$$U = \big\{(t, t_0, y_0) \;\big|\; |t - t_0| < a'/2,\ |t_0 - t^*| < a'/2,\ |y_0 - y^*| < b/2\big\}.$$

If F is smooth on D, then $(t, t_0, y_0) \mapsto y(t, t_0, y_0)$ is smooth on U.

REMARK. It is customary to summarize the result about continuity qualitatively by saying that the unique solution depends continuously on the initial data.

PROOF OF CONTINUITY. Let us first check that there is indeed a unique solution for each pair (t_0, y_0) in question and that its graph, as a function of t, lies in

$$R' = \{(t, y) \mid |t - t^*| \le a' \text{ and } |y - y^*| \le b\}.$$

For this purpose, fix t_0 and y_0 with $|t_0 - t^*| \le a'/2$ and $|y_0 - y^*| \le b/2$. Use of the triangle inequality shows that the closed set with $|t - t_0| < a'/2$ and $|y - y_0| < b/2$ lies within R. Thus $|F| \le M$ on this set. Theorem 4.1 shows that there exists a solution with graph in this smaller set for $|t - t_0| < a''$, where $a'' = \min\{a'/2, (b/2)/M\}$. Now

$$\min\{a'/2, b/(2M)\} = \tfrac{1}{2}\min\{a', b/M\} = \tfrac{1}{2}a',$$

and hence there exists a solution for $|t - t_0| < a'/2$ with graph in R. This solution $y(t, t_0, y_0)$ is unique by Theorem 4.2, and it is the result of the construction in the proof of Theorem 4.1.

The idea is to trace through the construction in the proof of Theorem 4.1 and to see that the function $(t, t_0, y_0) \mapsto y(t, t_0, y_0)$ is the uniform limit of explicit continuous functions on U. Imitating a part of the proof of Theorem 4.1, we define, for (t, t_0, y_0) in U,

$$\begin{aligned}
y_0(t, t_0, y_0) &= y_0, \\
y_1(t, t_0, y_0) &= y_0 + \int_{t_0}^{t} F(s, y_0(s, t_0, y_0))\, ds, \\
&\ \ \vdots \\
y_m(t, t_0, y_0) &= y_0 + \int_{t_0}^{t} F(s, y_{m-1}(s, t_0, y_0))\, ds.
\end{aligned}$$

We shall show by induction that $y_n(t, t_0, y_0)$ is continuous on U. Certainly $y_0(t, t_0, y_0)$ is continuous on U.

For the inductive step we need a preliminary calculation. Let I_1 be the closed interval between t_0 and t, and let I_2 be the closed interval between t_0' and t'. Suppose we have two functions f_1 and f_2 of a variable s such that

(i) f_1 is defined for s between t_0 and t with $|f_1| \le M$ there,
(ii) f_2 is defined for s between t_0' and t' with $|f_2| \le M$ there, and
(iii) $|f_1(s) - f_2(s)| \le \epsilon$ on their common domain.

If a' is $\ge$ the maximum distance among t_0, t, t_0', t', let us show that

$$\Big| \int_{t_0}^{t} f_1(s)\, ds - \int_{t_0'}^{t'} f_2(s)\, ds \Big| \le M(|t_0 - t_0'| + |t - t'|) + a'\epsilon. \qquad (*)$$

To show this for all possible order relations on the set $\{t_0, t_0', t, t'\}$, we observe that there is no loss of generality in assuming that t_0 is the smallest member of the set. There are then six cases.

Case 1. $t_0 \le t_0' \le t' \le t$, so that (iii) applies on $[t_0', t']$. Then

$$\int_{t_0}^{t} f_1(s)\,ds - \int_{t_0'}^{t'} f_2(s)\,ds = \int_{t_0}^{t_0'} f_1(s)\,ds + \int_{t_0'}^{t'} (f_1(s) - f_2(s))\,ds + \int_{t'}^{t} f_1(s)\,ds$$

and hence

$$\Big|\int_{t_0}^{t} f_1(s)\,ds - \int_{t_0'}^{t'} f_2(s)\,ds\Big| \le M|t_0' - t_0| + \epsilon|t' - t_0'| + M|t - t'|.$$

Therefore $(*)$ holds in this case.

Case 2. $t_0 \le t_0' \le t \le t'$, so that (iii) applies on $[t_0', t]$. Then

$$\int_{t_0}^{t} f_1(s)\,ds - \int_{t_0'}^{t'} f_2(s)\,ds = \int_{t_0}^{t_0'} f_1(s)\,ds + \int_{t_0'}^{t} (f_1(s) - f_2(s))\,ds - \int_{t}^{t'} f_2(s)\,ds,$$

and hence

$$\Big|\int_{t_0}^{t} f_1(s)\,ds - \int_{t_0'}^{t'} f_2(s)\,ds\Big| \le M|t_0' - t_0| + \epsilon|t - t_0'| + M|t' - t|.$$

Therefore $(*)$ holds in this case.

Case 3. $t_0 \le t \le t' \le t_0'$. Then

$$\Big|\int_{t_0}^{t} f_1(s)\,ds\Big| \le M|t - t_0| \le M(|t_0' - t_0| - |t_0' - t'|)$$

and

$$\Big|\int_{t_0'}^{t'} f_2(s)\,ds\Big| \le M|t_0' - t'|,$$

so that $(*)$ holds in this case.

Case 4. $t_0 \le t' \le t_0' \le t$. Then

$$\Big|\int_{t_0}^{t} f_1(s)\,ds\Big| \le M|t - t_0| = M(|t_0' - t_0| + |t - t_0'|)$$

and

$$\Big|\int_{t_0'}^{t'} f_2(s)\,ds\Big| \le M|t' - t_0'| = M(|t - t'| - |t - t_0'|),$$

so that $(*)$ holds in this case.

Case 5. $t_0 \le t \le t_0' \le t'$. Then

$$\Big| \int_{t_0}^{t} f_1(s)\,ds \Big| \le M|t - t_0| \le M|t_0' - t_0|$$

and
$$\Big| \int_{t_0'}^{t'} f_2(s)\,ds \Big| \le M|t' - t_0'| \le M|t' - t|,$$

so that $(*)$ holds in this case.

Case 6. $t_0 \le t' \le t \le t_0'$. Then

$$\Big| \int_{t_0}^{t} f_1(s)\,ds \Big| \le M|t - t_0| = M(|t_0' - t_0| - |t_0' - t|)$$

and
$$\Big| \int_{t_0'}^{t'} f_2(s)\,ds \Big| \le M|t_0' - t'| = M(|t_0' - t| + |t' - t|),$$

so that $(*)$ holds in this case.

With $(*)$ proved we can now proceed with the inductive step to show that $y_n(t, t_0, y_0)$ is continuous on U. Thus assume that $y_{n-1}(t, t_0, y_0)$ is continuous on U. If (t, t_0, y_0) and (t', t_0', y_0') are in U, then

$$\begin{aligned} &y_n(t, t_0, y_0) - y_n(t', t_0', y_0') \\ &\quad = (y_0 - y_0') + \int_{t_0}^{t} F(s, y_{n-1}(s, t_0, y_0))\,ds - \int_{t_0'}^{t'} F(s, y_{n-1}(s, t_0', y_0'))\,ds \\ &\quad = (y_0 - y_0') + \int_{t_0}^{t} f_1(s)\,ds - \int_{t_0'}^{t'} f_2(s)\,ds, \end{aligned}$$

where $f_1(s) = F(s, y_{n-1}(s, t_0, y_0))$ and $f_2(s) = F(s, y_{n-1}(s, t_0', y_0'))$. Thus $(*)$ gives

$$|y_n(t, t_0, y_0) - y_n(t', t_0', y_0')| \le |y_0 - y_0'| + M(|t_0 - t_0'| + |t - t'|) + a'\epsilon \quad (**)$$

if ϵ is chosen such that $|f_1(s) - f_2(s)| \le \epsilon$ on the common domain of f_1 and f_2.

Let $\epsilon > 0$ be given, and choose some $\delta > 0$ for uniform continuity of F on the set R. By uniform continuity of y_{n-1}, choose $\eta > 0$ such that

$$|y_{n-1}(s, t_0, y_0) - y_{n-1}(s, t_0', y_0')| < \delta \quad \text{whenever } |(s, t_0, y_0) - (s, t_0', y_0')| < \eta.$$

Then $|(s, t_0, y_0) - (s, t_0', y_0')| < \eta$ implies $|f_1(s) - f_2(s)| \le \epsilon$ on the common domain of f_1 and f_2, and hence $(**)$ holds. Therefore y_n is continuous as a function on U. This completes the induction.

We know that $y_n(t, t_0, y_0)$ converges to a solution $y(t, t_0, y_0)$ uniformly in t if (t_0, y_0) is fixed. Let us see that the convergence is in fact uniform in (t, t_0, y_0). The proof of Theorem 4.1 yielded the estimate

$$|y_n(t, t_0, y_0) - y_{n-1}(t, t_0, y_0)| \le \frac{M}{k} \frac{k^n (a')^n}{n!},$$

and this is independent of (t, t_0, y_0). Therefore the Weierstrass M test shows that $y_n(t, t_0, y_0)$ converges to $y(t, t_0, y_0)$ uniformly on U. The uniform limit of continuous functions is continuous by Proposition 2.21, and hence $y(t, t_0, y_0)$ is continuous.

PROOF OF SMOOTHNESS. Under the assumption that F is smooth on D, we are to prove that $y(t, t_0, y_0)$ is smooth on U. We return to the earlier proof of continuity of $y(t, t_0, y_0)$ and show that each $y_n(t, t_0, y_0)$ is smooth. This smoothness is trivial for $n = 0$, we assume inductively that $y_{n-1}(t, t_0, y_0)$ is smooth, and we form

$$y_n(t, t_0, y_0) = y_0 + \int_{t_0}^{t} F(s, y_{n-1}(s, t_0, y_0))\, ds.$$

The function on the right side is the composition of $(t, t_0, y_0) \mapsto (t, t_0, t_0, y_0)$ followed by $(t, t_0, s_0, y_0) \mapsto \int_{t_0}^{t} F(s, y_{n-1}(s, s_0, y_0))\, ds$. The chain rule (Theorem 3.10), the Fundamental Theorem of Calculus (Theorem 1.32), and Proposition 3.28 allow us to compute partial derivatives of this function, and another argument with $(*)$ allows us to see that the partial derivatives are continuous. There is no difficulty in iterating this argument, and we conclude that $y_n(t, t_0, y_0)$ is smooth.

The same argument in the proof of Theorem 4.1 that enabled us to estimate the size of $y_n(t, t_0, y_0) - y_{n-1}(t, t_0, y_0)$ allows us to estimate any iterated partial derivative of this difference. New constants enter the estimate, but the qualitative result is the same, namely that any iterated partial derivative of $y_n(t, t_0, y_0)$ converges uniformly to that same iterated partial derivative of $y(t, t_0, y_0)$. Applying Theorem 1.23, we see that $y(t, t_0, y_0)$ is smooth.

CONCLUDING REMARK. Sometimes a given system $y' = F(t, y)$ with initial condition $y(t_0) = y_0$ involves parameters in the definition of F, so that effectively the system is $y' = F(t, y, \lambda_1, \dots, \lambda_k)$. A natural problem is to find conditions under which the dependence of the solution on the k parameters is continuous or smooth. The answer is that this problem can be reduced to the problem addressed by Theorem 4.3. We simply introduce k additional variables z_j, one for each parameter λ_j, together with new equations $z_j' = 0$ and new initial conditions $z_j(t_0) = \lambda_j$.

4. Integral Curves

If U is an open subset of $\mathbb{R}^n$, then a **vector field** on U may be defined as a function $X : U \to \mathbb{R}^n$. The vector field is **smooth** if X is a smooth function. In classical notation, X is written $X = \sum_{j=1}^n a_j(x_1, \dots, x_n) \frac{\partial}{\partial x_j}$, and the function carries $(x_1, \dots, x_n)$ to $(a_1(x_1, \dots, x_n), \dots, a_n(x_1, \dots, x_n))$. The traditional geometric interpretation of X is to attach to each point p of U the vector $X(p)$ as an arrow based at p. This interpretation is appropriate, for example, if X represents the velocity vector at each point in space of a time-independent fluid flow.

We have defined the term "path" in a metric space to mean a continuous function from a closed bounded interval of $\mathbb{R}^1$ into the metric space. The term **curve** in a metric space is used to refer to a continuous function from an open interval of $\mathbb{R}^1$ into the metric space.

A standard problem in connection with vector fields on an open subset U of $\mathbb{R}^2$ is to try to draw curves within U with the property that the tangent vector to the curve at any point matches the arrow for the vector field. An illustration occurs in Figure 4.2. This section abstracts and generalizes this kind of curve.

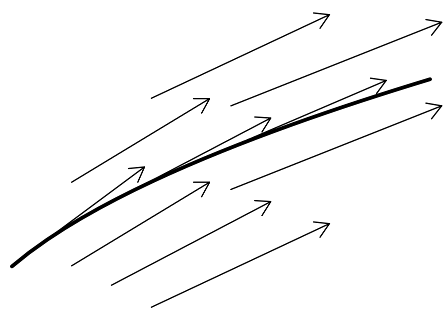


FIGURE 4.2. Integral curve of a vector field.

Let $X : U \to \mathbb{R}^n$ be a smooth vector field on U. A curve $c(t)$ is an **integral curve** for X if c is smooth and $c'(t) = X(c(t))$ for all t in the domain of c. Depending on one's interpretation of the informal wording in the previous paragraph, the present definition is perhaps more demanding than the definition given for $\mathbb{R}^2$ above: the expression $c'(t)$ involves both magnitude and direction, and the present definition insists that both ingredients match with $X(c(t))$, not just the direction.

Proposition 4.4. Let $X : U \to \mathbb{R}^n$ be a smooth vector field on an open subset U of $\mathbb{R}^n$, and let p be in U. Then there exist an $\varepsilon > 0$ and an integral curve $c : (-\varepsilon, \varepsilon) \to U$ such that $c(0) = p$. Any two integral curves c and d for X having $c(0) = d(0) = p$ coincide on the intersection of their domains.

PROOF. Apart from the smoothness the first conclusion is just a restatement of a special case of Theorem 4.1 in different notation. The conditions on c are that c be a solution of $c' = X(c)$ and that $c(0) = p$. The existence of a solution is immediate from Theorem 4.1 if we put $F = X$, $c = y$, $t_0 = 0$, and $y_0 = p$. The way in which this application of Theorem 4.1 is a special case and not the general case is that F is independent of t here. The smoothness of c follows from Theorem 4.3, and the uniqueness follows from Theorem 4.2.

The interest is not only in Proposition 4.4 in isolation but also in what happens to the integral curves when X is part of a family of vector fields.

Proposition 4.5. Let $X^{(1)}, \dots, X^{(m)}$ be smooth vector fields on an open subset U of $\mathbb{R}^n$, let p be in U, and let V be a bounded open neighborhood of 0 in $\mathbb{R}^m$. For λ in V, put $X_\lambda = \sum_{j=1}^m \lambda_j X^{(j)}$. Then there exist an $\varepsilon > 0$ and a system of integral curves $c(t, \lambda)$, defined for $t \in (-\varepsilon, \varepsilon)$ and $\lambda \in V$, such that $c(\,\cdot\,, \lambda)$ is an integral curve for X_λ with $c(0, \lambda) = p$. Each curve $c(t, \lambda)$ is unique, and the function $c : (-\varepsilon, \varepsilon) \times V \to U$ is smooth. If $m = n$, if the vectors $X^{(1)}(p), \dots, X^{(n)}(p)$ are linearly independent, and if δ is any positive number less than ε, then the Jacobian matrix of $\lambda \mapsto c(\delta, \lambda)$ at $\lambda = 0$ is nonsingular.

REMARK. In the final conclusion of this proposition, the open neighborhood of 0 within V is allowed to depend on δ. It follows from the final conclusion that the Inverse Function Theorem (Theorem 3.17) and its corollary (Corollary 3.21) are applicable to the mapping $\lambda \mapsto c(\delta, \lambda)$ at $\lambda = 0$. These results produce a smooth inverse function carrying an open subneighborhood of 0 within V onto an open subneighborhood of p of U. In effect the inverse function assigns locally defined coordinates in λ space to a neighborhood of U.

PROOF. We set up the system of equations $c' = X_\lambda \circ c$, i.e.,

$$c_i' = \sum_{j=1}^m \lambda_j X_i^{(j)}(c),$$

with initial condition $c(0) = p$. This is a smooth system of the kind considered in Theorem 4.3, and the λ_j with $1 \le j \le m$ are parameters. The parameters are handled by the concluding remark in Section 3: we obtain unique solutions $c(t, \lambda)$ for t in some open interval $(-\varepsilon, \varepsilon)$, and $(t, \lambda) \mapsto c(t, \lambda)$ is smooth.

Now suppose that $m = n$, that the vectors $X^{(1)}(p), \dots, X^{(n)}(p)$ are linearly independent, and that $0 < \delta < \varepsilon$. The function c satisfies

$$c_i'(t, \lambda) = \sum_{j=1}^n \lambda_j X_i^{(j)}(c(t, \lambda)), \tag{$*$}$$

and we use this information to compute the Jacobian matrix of $\lambda \mapsto c(\delta, \lambda)$ at $\lambda = 0$. The Fundamental Theorem of Calculus, Proposition 3.28, and $(*)$ give

$$\begin{aligned}\frac{\partial c_i}{\partial \lambda_j}(\delta, \lambda) &= \frac{\partial c_i}{\partial \lambda_j}(0, \lambda) + \int_0^\delta \frac{\partial c_i'}{\partial \lambda_j}(t, \lambda)\, dt \\ &= \frac{\partial c_i}{\partial \lambda_j}(0, \lambda) + \frac{\partial}{\partial \lambda_j}\int_0^\delta c_i'(t, \lambda)\, dt \\ &= \frac{\partial c_i}{\partial \lambda_j}(0, \lambda) + \int_0^\delta X_i^{(j)}(c(t, \lambda))\, dt + \sum_{k=1}^n \lambda_k \frac{\partial}{\partial \lambda_j}\int_0^\delta X_i^{(k)}(c(t, \lambda))\, dt.\end{aligned}$$

Now $c_i(0, \lambda) = p_i$ for all λ, and hence $\frac{\partial c_i}{\partial \lambda_j}(0, \lambda)\big|_{\lambda=0} = 0$. Also, $c(t, 0)$ is constant in t by $(*)$, and the constant is $c(0, 0) = p$. Finally when λ is set equal to 0 in the term $\sum_{k=1}^n \lambda_k \frac{\partial}{\partial \lambda_j}\int_0^\delta X_i^{(k)}(c(t, \lambda))\, dt$, each λ_k becomes 0, and thus the whole term becomes 0. Thus the above equation specializes at $\lambda = 0$ to

$$\frac{\partial c_i}{\partial \lambda_j}(\delta, \lambda)\Big|_{\lambda=0} = 0 + \delta X_i^{(j)}(p) + 0.$$

The vectors $X^{(j)}(p)$ are by assumption linearly independent, and hence the determinant of the matrix $[X_i^{(j)}(p)]$ is not 0. Consequently the Jacobian matrix $\lambda \mapsto c(\delta, \lambda)$ at $\lambda = 0$ is nonsingular if $\delta \neq 0$.

5. Linear Equations and Systems, Wronskian

Recall from Section 1 that a **linear ordinary differential equation** is defined to be an equation of the type

$$a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y' + a_0(t)y = q(t)$$

with real or complex coefficients. The equation is **homogeneous** if q is the 0 function, **inhomogeneous** in general. In order for the existence and uniqueness theorems of Section 1 to apply, we need to be able to solve for $y^{(n)}$ and have all coefficients be continuous afterward. Thus we assume that $a_n(t) = 1$ and that $a_{n-1}(t), \ldots, a_0(t)$ and $q(t)$ are continuous on some open interval.

Even in simple cases, the theory is helped by converting a single equation to a system of first-order equations. In Section 1 we saw an indication that a way to make this conversion is to put

$$\begin{aligned} y_1 &= y & \qquad\qquad && y_1' &= y_2 \\ y_2 &= y' & && y_2' &= y_3 \\ &\vdots & \text{and get} && &\vdots \\ y_{n-1} &= y^{(n-2)} & && y_{n-1}' &= y_n \\ y_n &= y^{(n-1)} & && y_n' &= -a_0(t)y_1 - \cdots - a_{n-1}y_n + q(t). \end{aligned}$$

If we change the meaning of the symbol y from a scalar-valued function to the vector-valued function $y = (y_1, \dots, y_n)$, then we arrive at the system

$$y' = A(t)y + Q(t),$$

where $A(t)$ is the n-by-n matrix of continuous functions given by

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0(t) & -a_1(t) & -a_2(t) & \cdots & -a_{n-1}(t) \end{pmatrix}$$

and $Q(t)$ is the n-component column vector of continuous functions given by

$$Q(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ q(t) \end{pmatrix}.$$

In a general linear first-order system of the kind we shall study, $A(t)$ can be any n-by-n matrix of continuous functions and $Q(t)$ can be any column vector of continuous functions; thus the first-order system obtained by conversion of a single n^{th}-order equation is of quite a special form among all first-order linear systems.

For a system $y' = A(t)y + Q(t)$ as above, the Lipschitz condition for the function $F(t, y) = A(t)y + Q(t)$ is automatic, since

$$|F(t, y) - F(t, y^*)| = |A(t)(y - y^*)| \le \|A(t)\| |y - y^*|$$

and since the function $t \mapsto \|A(t)\|$ is bounded on any compact subinterval of our domain interval. By the uniqueness theorem (Theorem 4.2), a unique solution to the system is determined by data (t_0, y_0), the local solution corresponding to (t_0, y_0) being the one satisfying the initial condition that the vector $y(t_0)$ equal the vector y_0. If we track down what these data correspond to in the case of a single n^{th}-order equation, we see that a unique solution to a single n^{th}-order equation of the kind described above is determined by initial values at a point t_0 for the scalar-valued solution and all its derivatives through order $n-1$.

First-order linear systems of size one can be solved explicitly in terms of known functions and integrations. Specifically the single homogeneous first-order equation $y' = a(t)y$ is solved by $y(t) = c\exp\big(\int^t a(s)\,ds\big)$, and the solution of a single inhomogeneous first-order equation can be reduced to the homogeneous case by the variation-of-parameters formula that appears later in this section. However, there need not be such an elementary solution of a first-order linear system of size two, not even a system that comes from a single second-order equation. Elementary solutions exist when the coefficient matrix has constants as entries, and we shall address that case in the next two sections. Sometimes one can write down tidy power-series solutions when the coefficient matrix has nonconstant entries, and we shall take up that matter later in the chapter. For now, we develop some general theory about first-order linear systems, beginning with the homogeneous case. The linearity implies that the set of solutions to the system $y' = A(t)y$ on an open interval is a vector space (of vector-valued functions) in the sense that it is closed under addition and scalar multiplication.

Theorem 4.6. Let $y' = A(t)y$ be a homogeneous linear first-order n-by-n system with $A(t)$ continuous for $a < t < b$. Then

(a) any solution on a subinterval (a', b') extends to a solution on the whole interval (a, b),

(b) the dimension of the vector space of solutions on any subinterval (a', b') is exactly n,

(c) if $v_1(t), \dots, v_r(t)$ are solutions on an interval (a', b') and if t_0 is in that interval, then $v_1, \dots, v_r$ are linearly independent functions if and only if the column vectors $v_1(t_0), \dots, v_n(t_0)$ are linearly independent.

PROOF. We begin by proving (c). If $c_1v_1(t) + \cdots + c_rv_r(t)$ is identically 0 for constants $c_1, \dots, c_r$ not all 0, then $c_1v_1(t_0) + \cdots + c_rv_r(t_0) = 0$ for the same constants. Conversely suppose that $c_1v_1(t_0) + \cdots + c_rv_r(t_0) = 0$ for constants not all 0. Put $v(t) = c_1v_1(t) + \cdots + c_rv_r(t)$. Then $v(t)$ and the 0 function are solutions of the system satisfying the same initial conditions—that they are 0 at t_0. By the uniqueness theorem (Theorem 4.2), $v(t)$ is the 0 function. This proves (c).

The upper bound in (b) is immediate from (c) since the dimension of the space of n-component column vectors is n.

Let us prove that n is a lower bound for the dimension in (b) if the interval containing t_0 is sufficiently small. By the existence theorem (Theorem 4.1), there exists a solution $v_j(t)$ on some interval $|t-t_0|<\varepsilon_j$ such that $v_j(t_0)=e_j$. The $v_j(t)$ are then solutions on $|t-t_0|<\varepsilon$ with $\varepsilon=\min\{\varepsilon_1,\dots,\varepsilon_n\}$, and they are linearly independent by (c). Hence the dimension of the space of solutions is at least n on the interval $|t-t_0|<\varepsilon$ or on any subinterval containing t_0.

We are not completely done with proving (b), but let us now prove (a). Let $v(t)$ be a solution on (a',b'). If we have a collection of solutions on different intervals containing (a',b') and each pair of solutions is consistent on their common domain, then the union of the solutions is a solution. Consequently we may assume that $v(t)$ does not extend to a solution on any larger interval. We are to prove that $(a',b')=(a,b)$. Suppose on the contrary that $b'<b$. We use $t_0=b'$ in the previous paragraph of the proof; the result is that on some interval $|t-b'|<\varepsilon$ with ε sufficiently small and at least small enough so that $a'<b'-\varepsilon$, the space of solutions has dimension n with a basis $\{v_1,\dots,v_n\}$. By (c), the column vectors $v_1(b'-\varepsilon),\dots,v_n(b'-\varepsilon)$ are linearly independent, and thus the restrictions of $v_1,\dots,v_n$ to $(b'-\varepsilon,\ b')$ are linearly independent. The restriction of $v(t)$ to the interval $(b'-\varepsilon,\ b')$ is a solution, and thus there exist constants $c_1,\dots,c_n$ such that

$$v(t)=c_1v_1(t)+\cdots+c_nv_n(t)\qquad \text{for } b'-\varepsilon<t<b'.$$

But then the function equal to $v(t)$ on (a',b') and equal to $c_1v_1(t)+\cdots+c_nv_n(t)$ on $(b'-\varepsilon,b'+\varepsilon)$ extends $v(t)$ to a solution on a larger interval and contradicts the maximality of the domain of $v(t)$. This proves that $b'=b$. Similarly we find that $a'=a$. This proves (a).

We return to the unproved part of (b). Fix t_0 in (a',b'). On a subinterval about t_0, the space of solutions has dimension n, as we have already proved. Let $\{v_1,\dots,v_n\}$ be a basis. By (a), we can extend $v_1,\dots,v_n$ to solutions on (a',b'). Then the space of solutions on (a',b') has dimension at least n, and (b) is now completely proved.

EXAMPLE. Let us illustrate the content of Theorem 4.6 by means of a single second-order equation, namely $y''+y=0$. We know that $c_1\cos t+c_2\sin t$ is a solution for every pair of constants c_1 and c_2. To convert the equation to a system, we introduce $y_1=y$ and $y_2=y'$. The system is then

$$\begin{aligned} y_1' &= y_2,\\ y_2' &= -y_1, \end{aligned}$$

and hence the matrix is $A(t)=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$, a matrix of constants. The scalar-valued solutions $\cos t$ and $\sin t$ of $y''+y=0$ correspond to the vector-valued solutions

$\begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}$ and $\begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$, respectively; each of these has a scalar-valued solution in its first entry and the derivative in the second entry. In either case, both solutions are defined on the interval $(-\infty, +\infty)$. The theorem says that the restrictions of these two functions to any subinterval span the solutions on that subinterval. According to (c), the linear independence of the scalar-valued solutions $\cos t$ and $\sin t$ is reflected by the linear independence of the column vectors $\begin{pmatrix} \cos t_0 \\ -\sin t_0 \end{pmatrix}$ and $\begin{pmatrix} \sin t_0 \\ \cos t_0 \end{pmatrix}$ for any t_0 in $(-\infty, +\infty)$. The latter independence we can see immediately by observing that the matrix $\begin{pmatrix} \cos t_0 & \sin t_0 \\ -\sin t_0 & \cos t_0 \end{pmatrix}$ has determinant equal to 1 and not 0.

The kind of matrix formed in the previous example is a useful tool when generalized to an arbitrary homogeneous linear system, and it has a customary name. Let $v_1(t), \dots, v_n(t)$ be solutions of an n-by-n homogeneous linear system $y' = A(t)y$ with $A(t)$ continuous. The **Wronskian matrix** of $v_1, \dots, v_n$ is the n-by-n matrix whose j^{th} column is v_j. If $v_{i,j}$ denotes the i^{th} entry of the j^{th} solution, then

$$W(t) = \begin{pmatrix} v_{1,1}(t) & \cdots & v_{1,n}(t) \\ \vdots & \ddots & \vdots \\ v_{n,1}(t) & \cdots & v_{n,n}(t) \end{pmatrix}.$$

Since each column of $W(t)$ is a solution, we obtain the matrix identity $W'(t) = A(t)W(t)$.

EXAMPLE, CONTINUED. In the case of the single second-order equation $y'' + y = 0$, we listed two linearly independent scalar-valued solutions as $\cos t$ and $\sin t$. When the equation is converted into a 2-by-2 homogeneous linear system, the Wronskian matrix is

$$W(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}.$$

For a general n^{th}-order equation with $v_1, \dots, v_n$ as scalar-valued solutions, the Wronskian matrix of the associated system is

$$W(t) = \begin{pmatrix} v_1(t) & \cdots & v_n(t) \\ v_1'(t) & \cdots & v_n'(t) \\ \vdots & \ddots & \vdots \\ v_1^{(n-1)}(t) & \cdots & v_n^{(n-1)}(t) \end{pmatrix}.$$

Proposition 4.7. If $v_1(t), \dots, v_n(t)$ are solutions on an interval of an n-by-n homogeneous linear system $y' = A(t)y$ with $A(t)$ continuous, then the following are equivalent:

(a) $v_1, \dots, v_n$ are linearly independent solutions,

(b) $\det W(t)$ is nowhere 0,
(c) $\det W(t)$ is somewhere nonzero.

PROOF. By Theorem 4.6c, (a) here is equivalent to the linear independence of $v_1(t_0), \ldots, v_n(t_0)$, no matter what t_0 we choose, hence is equivalent to the condition $\det W(t_0) \neq 0$, no matter what t_0 we choose. The proposition follows.

We shall use the Wronskian matrix of a homogeneous system to analyze the solutions of any corresponding inhomogeneous system.

Proposition 4.8. For an inhomogeneous linear system $y' = A(x)y + Q(t)$ with $A(t)$ and $Q(t)$ continuous for $a < t < b$, any solution $y^*(t)$ on a subinterval (a', b') of (a, b) extends to be a solution on (a, b), and the most general solution $y(t)$ is of the form $y(t) = h(t) + y^*(t)$, where $y^*(t)$ is one solution of $y' = A(t)y + Q(t)$ and $h(t)$ is an arbitrary solution of the homogeneous system $y' = A(t)y$.

PROOF. If y^* and y^{**} are two solutions of $y' = A(t)y + Q(t)$ on (a', b'), then $(y^{**} - y^*)'(t) = (A(t)y^{**}(t) + Q(t)) - (A(t)y^*(t) + Q(t)) = A(t)(y^* - y^{**})(t)$, and $h = y^{**} - y^*$ solves $y' = A(t)y$ on (a', b'). Conversely if h solves $y' = A(t)y + Q(t)$ on (a', b'), then

$$\begin{aligned}(y^* + h)'(t) &= y^{*\prime}(t) + h'(t)\\ &= (A(t)y^*(t) + Q(t)) + A(t)h(t) = A(t)(y^* + h)(t) + Q(t),\end{aligned}$$

and $y^* + h$ is a solution of $y' = A(t)y + Q(t)$ on (a', b').

We are left with showing that any solution y^* of $y' = A(t)y + Q(t)$ on (a', b') extends to a solution on (a, b). As in the proof of Theorem 4.6a, we can form unions of functions and thereby assume that y^* cannot be extended to be a solution on a larger interval. The claim is that $(a', b') = (a, b)$. Assuming the contrary, suppose, for example, that $b' < b$. By the existence theorem (Theorem 4.1), there exists a solution $y^{**}(t)$ of $y' = A(t)y + Q(t)$ for $|t - b'| < \varepsilon$ if ε is small enough. By the result of the previous paragraph, $y^*(t) = y^{**}(t) + h(t)$ on $(b' - \varepsilon, b')$ for a suitable choice of h that solves the homogeneous system $y' = A(t)y$ on $(b' - \varepsilon, b')$. Since $y^{**}(t)$ is given as a solution of $y' = A(t)y + Q(t)$ on $(b' - \varepsilon, b' + \varepsilon)$ and since, by Theorem 4.6a, $h(t)$ extends to a solution of $y' = A(t)y$ on $(b' - \varepsilon, b' + \varepsilon)$, we see that $y^{**}(t) + h(t)$ extends to a solution of $y' = A(t)y + Q(t)$ on $(b' - \varepsilon, b' + \varepsilon)$. Then the function equal to $y^*(t)$ on (a', b') and to $y^{**}(t) + h(t)$ on $(b' - \varepsilon,\ b' + \varepsilon)$ extends $y^*(t)$ to a solution of $y' = A(t)y + Q(t)$ on a larger interval, namely $(a', b' + \varepsilon)$. We obtain a contradiction and conclude that b' must have equaled b. Similarly a' must equal a. Thus every solution of $y' = A(t)y + Q(t)$ on a subinterval extends to all of (a, b), and the proof is complete.

Theorem 4.9 (variation of parameters). For an inhomogeneous linear system $y' = A(x)y + Q(t)$ with $A(t)$ and $Q(t)$ continuous for $a < t < b$, let $v_1, \dots, v_n$ be linearly independent solutions of $y' = A(t)y$ on (a, b), and let $W(t)$ be their Wronskian matrix. Then a particular solution y^* of $y' = A(t)y + Q(t)$ on (a, b) is given by

$$y^*(t) = W(t)u(t), \qquad \text{where} \quad W(t)u'(t) = Q(t).$$

That is,

$$y^*(t) = W(t)\int^t W(s)^{-1}Q(s)\,ds.$$

REMARKS. Linearly independent solutions $v_1, \dots, v_n$ as in the statement exist by Theorem 4.6.

PROOF. For any differentiable vector-valued function $u(t)$, $y^*(t) = W(t)u(t)$ has

$$(y^*)' = W'u + Wu' = AWu + Wu' = Ay^* + Wu'.$$

Thus y^* will have $(y^*)' = Ay^* + Q$ if and only if $Wu' = Q$. Since Proposition 4.7 shows that $W(t)^{-1}$ exists and is continuous, we can solve $Wu' = Q$ for u.

EXAMPLE, CONTINUED. Now consider the single second-order inhomogeneous linear equation $y'' + y = \tan t$ on the interval $|t| < \pi/2$. We saw that we can take $W(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$. We set up the system

$$\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}\begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} 0 \\ \tan t \end{pmatrix}$$

of algebraic linear equations and solve for u_1' and u_2':

$$\begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}\begin{pmatrix} 0 \\ \tan t \end{pmatrix} = \begin{pmatrix} -\dfrac{\sin^2 t}{\cos t} \\ \sin t \end{pmatrix}.$$

A vector-valued function with derivative $\begin{pmatrix} u_1' \\ u_2' \end{pmatrix}$ for $|t| < \pi/2$ is

$$\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = \begin{pmatrix} \sin t - \log(1 + \sin t) + \log\cos t \\ -\cos t \end{pmatrix},$$

and we thus take $y^*(t) = (\cos t)u_1(t) + (\sin t)u_2(t)$. The most general solution of the given inhomogeneous equation is therefore $y^*(t) + c_1\cos t + c_2\sin t$.

6. Homogeneous Equations with Constant Coefficients

In this section and the next, we discuss first-order homogeneous linear systems with constant coefficients. The system is of the form $y' = Ay$ with A a matrix of constants. A single homogeneous n^{th}-order linear equation with constant coefficients can be converted into such a first-order system and can therefore be handled by the method applicable to all first-order homogeneous linear systems with constant coefficients. But such an equation can be handled more simply in a direct fashion, and we therefore isolate in this section the case of a single n^{th}-order equation. This section and the next will make use of material on polynomials from Section A8 of the appendix.

The equation to be studied in this section is of the form

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0$$

with coefficients in $\mathbb{C}$. Let us write this equation as $L(y) = 0$ for a suitable linear operator L defined on functions y of class C^n:

$$L = \Big(\frac{d}{dt}\Big)^n + a_{n-1}\Big(\frac{d}{dt}\Big)^{n-1} + \cdots + a_1\Big(\frac{d}{dt}\Big) + a_0.$$

The term a_0 is understood to act as a_0 times the identity operator. Since $\frac{d}{dt}\, e^{rt} = re^{rt}$, we immediately obtain

$$L(e^{rt}) = (r^n + a_{n-1}r^{n-1} + \cdots + a_1 r + a_0)e^{rt}.$$

The polynomial

$$P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

is called the **characteristic polynomial** of the equation, and the formula $L(e^{rt}) = P(r)e^{rt}$ shows that $y(t) = e^{rt}$ is a solution of $L(y) = 0$ if and only if r is a root of the characteristic polynomial. From Section A8 of the appendix, we know that the polynomial $P(\lambda)$ factors into the product of linear factors $\lambda - r$, the factors being unique apart from their order. Let us list the distinct roots, i.e., the distinct such complex numbers r, as $r_1, \dots, r_k$ with $k \le n$, and let us write m_j for the number of times that $\lambda - r_j$ occurs as a factor of $P(\lambda)$, i.e., the multiplicity of r_j as a root of P. Then we have $\sum_{j=1}^k m_j = n$ and

$$P(\lambda) = \prod_{j=1}^k (\lambda - r_j)^{m_j}.$$

Corresponding to this factorization of P is a factorization of L as

$$L = \prod_{j=1}^k \Big(\frac{d}{dt} - r_j\Big)^{m_j}.$$

On the right side the individual factors commute with each other because differentiation commutes with itself and with multiplication by constants. The following lemma therefore produces n solutions of the given equation $L(y) = 0$.

Lemma 4.10. For $m \geq 1$ and r in $\mathbb{C}$, all the functions $e^{rt}, te^{rt}, \ldots, t^{m-1}e^{rt}$ are solutions of the m^{th}-order differential equation

$$\Big(\frac{d}{dt} - r\Big)^m (y) = 0.$$

PROOF. Direct computation gives $\big(\frac{d}{dt} - r\big)(t^k e^{rt}) = kt^{k-1}e^{rt}$, and hence $\big(\frac{d}{dt} - r\big)^m (t^k e^{rt}) = k(k-1)\cdots(k-m+1)t^{k-m}e^{rt}$. The right side is 0 if $0 \leq k \leq m-1$, and the lemma follows.

Lemma 4.11. Let $r_1, \ldots, r_N$ be distinct complex numbers, and let m_j be N integers ≥ 1. Then the $\sum_{j=1}^N m_j$ functions

$$e^{r_j t}, te^{r_j t}, \ldots, t^{m_j - 1}e^{r_j t}, \qquad 1 \leq j \leq N,$$

are linearly independent over $\mathbb{C}$.

PROOF. Let $k \geq 1$ be an integer, let r be a complex number, and let $P(t)$ be a polynomial of degree $\leq k-1$. We allow $P(t)$ to be the 0 polynomial. Then

$$\tfrac{d}{dt}[(t^k + P(t))e^{rt}] = r(t^k + P(t))e^{rt} + ((k-1)t^{k-1} + P'(t))e^{rt},$$

from which it follows that

$$\tfrac{d}{dt}[(t^k + P(t))e^{rt}] = (rt^k + Q(t))e^{rt} \qquad (*)$$

with $Q(t)$ a polynomial of degree $\leq k-1$ or the 0 polynomial.

We shall prove by induction on N that if $P_1, \ldots, P_N$ are polynomials with complex coefficients such that $\sum_{j=1}^N P_j(t)e^{r_j t}$ is the 0 function, then all the P_j are 0 polynomials. For $N = 1$, if $P(t)e^{rt}$ is the 0 function, then $P(t)$ is the 0 function. Since a polynomial of degree $k \geq 0$ has at most k roots, we conclude that P has all coefficients 0. This disposes of the assertion for $N = 1$. Assume the result for $N-1$, and suppose that we are given that $\sum_{j=1}^{N-1} P_j(t)e^{r_j t} + P_N(t)e^{r_N t}$ is the 0 function, where $\{r_1, \ldots, r_{N-1}, r_N\}$ are distinct. Then

$$\sum_{j=1}^{N-1} P_j(t)e^{q_j t} + P_N(t) \qquad (**)$$

is the 0 function when $q_j = r_j - r_N$ for $j \leq N-1$. If P_N is the 0 polynomial, the inductive hypothesis shows that all P_j with $j \leq N-1$ are 0 polynomials. Otherwise let P_N have degree d, and differentiate $(**)$ $d+1$ times. If $P_j(t)$ for $j \leq N-1$ is the sum of $a_{n_j}t^{n_j}$ plus lower-degree terms, then $(*)$ shows that the result of the differentiation is that

$$\sum_{j=1}^{N-1} \big(a_{n_j}(q_j)^{d+1}t^{n_j} + \text{lower-degree terms}\big)e^{q_j t}$$

is the 0 function. By the inductive hypothesis each a_{n_j} has to be 0, and hence all coefficients of each P_j have to be 0 for $j \le N-1$. Then $P_N(t)$ is identically 0 and must be the 0 polynomial. This completes the induction.

If we are given a linear combination of the functions in the statement of the lemma that equals the 0 function, then we obtain a relation of the form $\sum_{j=1}^{N} P_j(t)e^{r_j t} = 0$, and we have just seen that this relation forces all P_j to be 0 polynomials. This completes the proof.

Proposition 4.12. Let the differential equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0,$$

with complex coefficients, have characteristic polynomial given by $P(\lambda) = \prod_{j=1}^{k}(\lambda - r_j)^{m_j}$ with $r_1, \dots, r_k$ distinct complex numbers and with the m_j integers ≥ 0 such that $\sum_{j=1}^{k} m_j = n$. Then the n functions

$$e^{r_j t}, te^{r_j t}, \dots, t^{m_j - 1}e^{r_j t}, \qquad 1 \le j \le k,$$

form a basis over $\mathbb{C}$ of the space of solutions of the given equation on any interval.

PROOF. Lemma 4.10 shows that the functions in question are solutions, Lemma 4.11 shows that they are linearly independent, and Theorem 4.6 shows that the dimension of the space of solutions on any interval is n. Since n linearly independent solutions have been exhibited, they must form a basis of the space of solutions.

If the equation in Proposition 4.12 happens to have real coefficients, it is meaningful to ask for a basis over $\mathbb{R}$ of the space of real-valued solutions. Since the coefficients are real, we have $L(\bar{y}) = \overline{L(y)}$ for all complex-valued functions y of class C^n, and it follows that the complex conjugate of any complex-valued solution is again a solution. Thus the real and imaginary parts of any complex-valued solution are real-valued solutions. Meanwhile, the characteristic polynomial P of the equation has real coefficients, and it follows that the set of roots of P is closed under complex conjugation. In addition, the multiplicity of a root equals the multiplicity of its complex conjugate. For any integer $k \ge 0$ and complex number $a + bi$ with $b \ne 0$, we have

$$\mathbb{C}t^k e^{(a+bi)t} + \mathbb{C}t^k e^{(a-bi)t} = \mathbb{C}t^k e^{at}\cos bt + \mathbb{C}t^k e^{at}\sin bt.$$

Thus $t^k e^{at}\cos bt$ and $t^k e^{at}\sin bt$ form a basis over $\mathbb{C}$ of the space spanned by $t^k e^{(a+bi)t}$ and $t^k e^{(a-bi)t}$. The functions $t^k e^{at}\cos bt$ and $t^k e^{at}\sin bt$ are real-valued, and thus we obtain a basis over $\mathbb{C}$ consisting of the real-valued solutions of the given equation if we retain the solutions $t^k e^{rt}$ with r real and we replace any pair $t^k e^{(a+bi)t}$ and $t^k e^{(a-bi)t}$ of solutions, $b \neq 0$, by the pair $t^k e^{at} \cos bt$ and $t^k e^{at} \sin bt$.

Let us see that these resulting functions form a basis over $\mathbb{R}$ of the *real* vector space of real-valued solutions. In fact, we know that they are linearly independent over $\mathbb{R}$ because they are linearly independent over $\mathbb{C}$. To see that they span, we take any real-valued solution and expand it as a complex linear combination of these functions. The imaginary part of this expansion exhibits 0 as a linear combination of the given functions, and the coefficients must be 0 by linear independence. Thus the constructed functions form a basis over $\mathbb{R}$ of the space of real-valued solutions.

7. Homogeneous Systems with Constant Coefficients

Having discussed linear homogeneous equations with constant coefficients, let us pass to the more general case of first-order homogeneous linear systems with constant coefficients. We write the system as $y' = Ay$ with A an n-by-n matrix of constants. In principle we can solve the system immediately. Namely, Proposition 3.13c tells us that $\frac{d}{dt}(e^{tA}) = Ae^{tA}$, so that each of the n columns of e^{tA} is a solution of $y' = Ay$. At $t = 0$, e^{tA} reduces to the identity matrix, and thus these n solutions are linearly independent at $t = 0$. By Theorem 4.6 these n solutions form a basis of all solutions on any subinterval (a, b) of $(-\infty, +\infty)$. The solution satisfying the initial condition $y(t_0) = y_0$ is $y(t) = e^{tA}e^{-t_0 A}y_0$, which is the particular linear combination $\sum_{j=1}^n c_j e^{tA} e_j$ of the columns of e^{tA} in which c_j is the number $c_j = (e^{-t_0 A} y_0)_j$.

In practice it is not so obvious how to compute e^{tA} except in special cases in which the exponential series can be summed entry by entry. Let us write down three model cases of this kind, and ultimately we shall see that we can handle general A by working suitably with these cases.

MODEL CASES.

(1) Let

$$C = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ & 0 & 1 & 0 & \cdots & 0 & 0 \\ & & 0 & 1 & \cdots & 0 & 0 \\ & & & \ddots & \ddots & \vdots & \vdots \\ & & & & 0 & 1 & 0 \\ & & & & & 0 & 1 \\ & & & & & & 0 \end{pmatrix}$$

be of size m-by-m with 0's below the main diagonal. Raising C to powers, we see that the $(i, j)^{\text{th}}$ entry of A^k is 1 if $j = i + k$ and is 0 otherwise. Hence

$$e^{tC} = \begin{pmatrix} 0 & t & \frac{1}{2!}t^2 & \frac{1}{3!}t^3 & \cdots & \frac{1}{(m-2)!}t^{m-2} & \frac{1}{(m-1)!}t^{m-1} \\ & 0 & t & \frac{1}{2!}t^2 & \cdots & \frac{1}{(m-3)!}t^{m-3} & \frac{1}{(m-2)!}t^{m-2} \\ & & 0 & t & \cdots & \vdots & \frac{1}{(m-3)!}t^{m-3} \\ & & & \ddots & \ddots & & \vdots \\ & & & & 0 & t & \frac{1}{2!}t^2 \\ & & & & & 0 & t \\ & & & & & & 0 \end{pmatrix}$$

with 0's below the main diagonal.

(2) Let

$$A = \begin{pmatrix} a & 1 & 0 & 0 & \cdots & 0 & 0 \\ & a & 1 & 0 & \cdots & 0 & 0 \\ & & a & 1 & \cdots & 0 & 0 \\ & & & \ddots & \ddots & \vdots & \vdots \\ & & & & a & 1 & 0 \\ & & & & & a & 1 \\ & & & & & & a \end{pmatrix},$$

so that $A = a1 + C$ with C as in the previous case. Since $a1$ and C commute, Proposition 3.13a shows that $e^{tA} = e^{at}e^{tC}$. In other words, e^{tA} is obtained by multiplying every entry of the matrix e^{tC} in the previous case by e^{at}. A matrix of this form A for some complex constant a and for some size m is said to be a **Jordan block**. Thus we know how to form e^{tA} if A is a Jordan block.

(3) Let A be block diagonal with each block being a Jordan block:

$$A = \begin{pmatrix} \text{block \#1} & & & \\ & \text{block \#2} & & \\ & & \ddots & \\ & & & \text{block \#}k \end{pmatrix}.$$

Then

$$e^{tA} = \begin{pmatrix} e^{t\,\text{block \#1}} & & & \\ & e^{t\,\text{block \#2}} & & \\ & & \ddots & \\ & & & e^{t\,\text{block \#}k} \end{pmatrix}.$$

Thus we know how to form e^{tA} if A is block diagonal with each block being a Jordan block. A matrix A of this kind is said to be in **Jordan form**.

The theorem reduces any computation of a matrix e^{tA} to this case.

Theorem 4.13 (Jordan normal form). For any square matrix A with complex entries, there exists a nonsingular complex matrix B such that $B^{-1}AB = J$ is in Jordan form.

REMARKS. This theorem comes from linear algebra, but knowledge of it is beyond the algebra prerequisites for this book. The proof is long and is not in the spirit of this text, and we shall omit it; however, the interested reader can find a proof in many linear algebra books. As a practical matter, the proof will not give us any additional information, since we already know that e^{tA} yields the solutions to $y' = Ay$ and the only remaining question is to convert the statement of the theorem into an explicit method of computation.

Let us see what Theorem 4.13 accomplishes. The solution of $y' = Ay$ with $y(t_0) = y_0$ is $y(t) = e^{(t-t_0)A}y_0$. Write $B^{-1}AB = J$ as in the proposition. Then Proposition 3.13d gives

$$\begin{aligned} y(t) = e^{(t-t_0)A}y_0 &= B(B^{-1}e^{(t-t_0)A}B)B^{-1}y_0 \\ &= Be^{(t-t_0)B^{-1}AB}B^{-1}y_0 = Be^{(t-t_0)J}B^{-1}y_0. \end{aligned}$$

If we can compute J, then Model Case 3 above tells us what $e^{(t-t_0)J}$ is. If we can compute B also, then we recover $y(t)$ explicitly.

The practical effect is that Theorem 4.13 gives us a method for calculating solutions. The idea behind the method is that the qualitative properties of B and J forced by the theorem are enough to lead us to explicit values of B and J. Let us go through the steps. A concrete example of J is

$$J = \begin{pmatrix} a & 1 & 0 & & & & & & & \\ 0 & a & 1 & & & & & & & \\ 0 & 0 & a & & & & & & & \\ & & & a & 1 & & & & & \\ & & & 0 & a & & & & & \\ & & & & & a & 1 & & & \\ & & & & & 0 & a & & & \\ & & & & & & & a & & \\ & & & & & & & & b & \\ & & & & & & & & & \ddots \end{pmatrix}.$$

It is helpful to know the extent of uniqueness in Theorem 4.13. The matrix J is actually unique up to permuting the order of the Jordan blocks. The matrix B is not at all unique but results from finding bases of certain subspaces of $\mathbb{C}^n$. The first step is to form the **characteristic polynomial**[2] $P(\lambda) = \det(\lambda 1 - A)$ of A. We have

$$\begin{aligned}\det(\lambda 1 - J) &= \det(\lambda 1 - B^{-1}AB) = \det(B^{-1}(\lambda 1 - A)B)\\ &= \det(B)^{-1}\det(\lambda 1 - A)\det(B) = \det(\lambda 1 - A),\end{aligned}$$

and thus J has the same characteristic polynomial as A. The characteristic polynomial of J is just the product of expressions $\lambda - d$ as d runs through the diagonal entries of J. According to Section A8 of the appendix, the factorization of a polynomial with complex coefficients and with leading coefficient 1 into first-degree expressions $\lambda - c$ is unique up to order, and thus the factorization of $P(\lambda)$ tells us the diagonal entries of J. We still need to know the sizes of the individual Jordan blocks.

The sizes of the Jordan blocks come from computing dimensions of various null spaces—or kernels, in the terminology of linear functions. If a occurs as a diagonal entry of J, think of forming $J - a1$ and its powers, and consider the dimension of the kernel of each power. For example, with the explicit matrix J that is written above, we have

$$J - a1 = \begin{pmatrix} 0 & 1 & 0 & & & & & & \\ 0 & 0 & 1 & & & & & & \\ 0 & 0 & 0 & & & & & & \\ & & & 0 & 1 & & & & \\ & & & 0 & 0 & & & & \\ & & & & & 0 & 1 & & \\ & & & & & 0 & 0 & & \\ & & & & & & & 0 & \\ & & & & & & & & \boxed{\text{nonsingular}} \end{pmatrix},$$

and $\dim\ker(J - a1)$ is the number of Jordan blocks of size ≥ 1 with a on the diagonal, namely 4 in this case. Next we consider $(J - a1)^2$. In this case,

$$(J - a1)^2 = \begin{pmatrix} 0 & 0 & 1 & & & & & & \\ 0 & 0 & 0 & & & & & & \\ 0 & 0 & 0 & & & & & & \\ & & & 0 & 0 & & & & \\ & & & 0 & 0 & & & & \\ & & & & & 0 & 0 & & \\ & & & & & 0 & 0 & & \\ & & & & & & & 0 & \\ & & & & & & & & \boxed{\text{nonsingular}} \end{pmatrix},$$

[2]Many books write the characteristic polynomial as $\det(A - \lambda 1)$, which is the same as the present polynomial if n is even but is its negative if n is odd. The present notation has the advantage that the notions of characteristic polynomial here and in the previous section coincide when an n^{th}-order equation is converted into a first-order system.

and $\dim\ker(J - a1)^2 = 7$. This number arises as the sum of the previous number and the number of Jordan blocks of size ≥ 2 with a on the diagonal. Thus $\dim\ker(J - a1)^2 - \dim\ker(J - a1)$ in general is the number of Jordan blocks of size ≥ 2 with a on the diagonal. Finally we consider $(J - a1)^3$. In this case, the upper left part of $(J - a1)^3$ corresponding to diagonal entry a is all 0, and the lower right part is nonsingular; hence $\dim\ker(J - a1)^3 = 8$. This number arises as the sum of the previous number and the number of Jordan blocks of size ≥ 3 with a on the diagonal. Thus in general, $\dim\ker(J - a1)^3 - \dim\ker(J - a1)^2$ is the number of Jordan blocks of size ≥ 3 with a on the diagonal. In our example, the number $\dim\ker(J - a1)^k$ remains at 8 for all $k \geq 3$ because 8 is the multiplicity of a as a root of $P(\lambda)$, and we are therefore done with diagonal entry a; our computation has shown that the numbers of Jordan blocks of sizes $1, 2, 3, 4, \dots,$ are $1, 2, 1, 0, \dots,$ and a check on the computation is that $1(1) + 2(2) + 3(1) = 8$.

Of course, we do not have J at our disposal for these calculations, but A yields the same numbers. In fact, we have $B(J - a1)^k B^{-1} = (A - a1)^k$, from which we see that $x \in \ker(A - a1)^k$ if and only if $B^{-1}x \in \ker(J - a1)k$. Hence

$$B(\ker(J - a1)^k) = \ker(A - a1)^k.$$

Since B is nonsingular, the dimension of the kernel of $(J - a1)^k$ equals the dimension of the kernel of $(A - a1)^k$. Consequently

$$\begin{aligned}
\dim\ker(A - a1) &= \#\{\text{Jordan blocks of size} \geq 1 \text{ with } a \text{ on diagonal}\}, \\
\dim\ker(A - a1)^2 - \dim\ker(A - a1)& \\
&= \#\{\text{Jordan blocks of size} \geq 2 \text{ with } a \text{ on diagonal}\}, \\
\dim\ker(A - a1)^3 - \dim\ker(A - a1)^2& \\
&= \#\{\text{Jordan blocks of size} \geq 3 \text{ with } a \text{ on diagonal}\}, \\
&\text{etc.}
\end{aligned}$$

Repeating this argument with the other roots of $P(\lambda)$, we find that we can determine J completely.

Calculating B requires working with vectors rather than dimensions. The columns of B are just $Be_1, \dots, Be_n$, and we seek a way of finding these. Fix attention on a root a of $P(\lambda)$. Consider an index i with $1 \leq i \leq n$, and suppose that the diagonal entry of J in column i is a. From the form of J, we see that either the i^{th} column of $J - a1$ is 0 or else it is e_{i-1}. In the latter case, index $i - 1$ corresponds to the same Jordan block. Using the identity $(A - a1)B = B(J - a1)$, we see that either

$$(A - a1)(Be_i) = B(J - a1)e_i = 0$$

or
$$(A - a1)(Be_i) = B(J - a1)e_i = Be_{i-1},$$

and index $i-1$ corresponds to the same Jordan block as index i in the latter case. Thus the vectors Be_i corresponding to the columns with diagonal entry a and with smallest index for a Jordan block lie in $\ker(A-a1)$. They are linearly independent since B is nonsingular, and the number of them is the number of Jordan blocks corresponding to diagonal entry a. We saw that this number equals $\dim\ker(A-a1)$. Hence the vectors Be_i corresponding to the smallest indices going with each Jordan block form a basis of $\ker(A-a1)$.

Similarly

$$(A-a1)^2(Be_i) = B(J-a1)^2 e_i = 0$$

or

$$(A-a1)^2(Be_i) = B(J-a1)^2 e_i = Be_{i-2},$$

and index $i-2$ corresponds to the same Jordan block as index i in the latter case. Thus the vectors Be_i corresponding to the columns with diagonal entry a and with smallest or next smallest index for a Jordan block lie in $\ker(A-a1)^2$. They are linearly independent since B is nonsingular, and the number of them is the sum of the previously computed number, namely $\dim\ker(A-a1)$, plus the number of Jordan blocks of size ≥ 2 that correspond to diagonal entry a. We saw that this sum equals $\dim\ker(A-a1)^2$. Hence the vectors Be_i corresponding to the two smallest indices going with each Jordan block form a basis of $\ker(A-a1)^2$. The new vectors Be_i are therefore vectors that we adjoin to a basis of $\ker(A-a1)$ to obtain a basis of $\ker(A-a1)^2$.

In setting up these vectors properly, however, we have to correlate the indices studied at the previous step with those being studied now. The relevant formula is that the new indices i have the property $(A-a1)Be_i = Be_{i-1}$. To obtain vectors with this consistency property, we would take a basis S_1 of $\ker(A-a1)$, extend it to a basis S_2 of $\ker(A-a1)^2$, discard the members of S_1, apply $A-a1$ to the members of S_2-S_1, and extend $(A-a1)(S_2-S_1)$ to a basis T_1 of $\ker(A-a1)$. Then $S_2' = (S_2-S_1)\cup T_1$ is a new basis of $\ker(A-a1)^2$.

We can continue the argument in this way. It is perhaps helpful to read the general discussion of the argument side by side with the explicit example that appears below. We continue to find that the construction of new basis vectors gets in the way of the necessary consistency property with the earlier basis vectors. Thus we really must start with the largest index k such that $\ker(A-a1)^k \neq \ker(A-a1)^{k-1}$. We extend a basis S_{k-1} of $\ker(A-a1)^{k-1}$ to a basis S_k of $\ker(A-a1)^k$, and form

$$(S_k - S_{k-1}) \cup (A-a1)(S_k-S_{k-1}) \cup \cdots \cup (A-a1)^{k-1}(S_k-S_{k-1}).$$

These vectors will be the columns of B corresponding to the largest Jordan blocks with diagonal entry a. The vectors in

$$(A-a1)^2(S_k-S_{k-1}) \cup \cdots \cup (A-a1)^{k-1}(S_k-S_{k-1})$$

are linearly independent in $\ker(A-a1)^{k-2}$; we extend this set to a basis S'_{k-2} of $\ker(A-a1)^{k-2}$, and we extend $S'_{k-2}\cup(A-a1)(S_k-S_{k-1})$ to a basis S'_{k-1} of $\ker(A-a1)^{k-1}$. The adjoined vectors, together with the result of applying powers of $A-a1$ to them, will be the columns of B corresponding to the next largest Jordan blocks with diagonal entry a. The process continues until we obtain a basis of $\ker(A-a1)^k$ with the necessary consistency property throughout. Then we repeat the process for the other roots of $P(\lambda)$ and assemble the result.

EXAMPLE. Let
$$A=\begin{pmatrix}4&1&-1\\-8&-2&2\\8&2&-2\end{pmatrix}.$$
The characteristic polynomial is $P(\lambda)=\det(\lambda 1-A)=\lambda^3$, whose factorization is evidently $P(\lambda)=(\lambda-0)^3$. Computing the kernel of A, we find that $\dim\ker A=2$, so that there are 2 Jordan blocks. Also, $A^2=0$, so that $\dim\ker A^2=3$ and the number of blocks of size ≥ 2 is $3-2=1$. Thus
$$J=\begin{pmatrix}0&1&0\\0&0&0\\0&0&0\end{pmatrix}.$$
We form a basis of $\ker A$ by solving $A\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=0$. The standard method of row reduction gives $x_1=-\frac{1}{4}x_2+\frac{1}{4}x_3$ with x_2 and x_3 arbitrary, so that a basis of $\ker A$ consists of $\begin{pmatrix}-\frac{1}{4}\\1\\0\end{pmatrix}$ and $\begin{pmatrix}\frac{1}{4}\\0\\1\end{pmatrix}$. We extend this to a basis of $\ker A^2=\mathbb{C}^3$ by adjoining, for example, the vector $v_1=\begin{pmatrix}1\\0\\0\end{pmatrix}$. Then $Av_1=\begin{pmatrix}4\\-8\\8\end{pmatrix}$. The vector Av_1 is in $\ker A$, and we extend it to a basis of $\ker A$ by adjoining, for example, $v_2=\begin{pmatrix}-1\\4\\0\end{pmatrix}$. Then v_1, Av_1, v_2 form a basis of $\ker A^2=\mathbb{C}^3$, and the above general method asks that these vectors be listed in the order Av_1, v_1, v_2. The matrix B is obtained by lining these vectors up as columns:
$$B=\begin{pmatrix}4&1&-1\\-8&0&4\\8&0&0\end{pmatrix}.$$
The result is easy to check. Computation shows that $B^{-1}=\begin{pmatrix}0&0&\frac{1}{8}\\1&\frac{1}{4}&-\frac{1}{4}\\0&\frac{1}{4}&\frac{1}{4}\end{pmatrix}$, and then one can carry out the multiplications to verify that $B^{-1}AB=J$.

8. Series Solutions in the Second-Order Linear Case

In this section we shall consider, in some detail, series solutions for two kinds of ordinary differential equations.

The first kind is

$$y'' + P(t)y' + Q(t)y = 0,$$

where $P(t)$ and $Q(t)$ are given by convergent power-series expansions for $|t| < R$:

$$\begin{aligned} P(t) &= a_0 + a_1 t + a_2 t^2 + \cdots, \\ Q(t) &= b_0 + b_1 t + b_2 t^2 + \cdots. \end{aligned}$$

We seek power-series solutions of the form

$$y(t) = c_0 + c_1 t + c_2 t^2 + \cdots.$$

The same methods and theorem that handle this first kind of equation apply also to n^{th}-order homogeneous linear equations and to first-order homogeneous systems when the leading coefficient is 1 and the other coefficients are given by convergent power series. The second-order case, however, is by far the most important for applications and is sufficiently illustrative that we shall limit our attention to it.

The idea in finding the solutions is to assume that we have a convergent power-series solution $y(t)$ as above, to substitute the series into the equation, and to sort out the conditions that are imposed on the unknown coefficients. Our theorems on power series in Section I.7 guarantee us that the operations of differentiation and multiplication of power series maintain convergence, and thus the result of substituting into the equation is that we obtain an equality of a convergent power series with 0. Corollary 1.39 then shows that all the coefficients of this last power series must be 0, and we obtain recursive equations for the unknown coefficients. There is one theorem about the equations under study, and it tells us that the power series for $y(t)$ that we obtain by these manipulations is indeed convergent; we state and prove this theorem shortly.

Let us go through the steps of finding the solutions. These steps turn out to be clearer when done in complete generality than when done for an example. Thus we shall first make the computation in complete generality, then state and prove the theorem, and finally consider an important example. The expansions of $y(t)$ and its derivatives are

$$\begin{aligned} y(t) &= c_0 + c_1 t + c_2 t^2 + \cdots, \\ y'(t) &= c_1 + 2c_2 t + 3c_3 t^2 + \cdots, \\ y''(t) &= 2c_2 + 3 \cdot 2c_3 t + 4 \cdot 3c_4 t^2 + \cdots. \end{aligned}$$

Substituting all the series into the given equation yields

$$\begin{aligned}(2\cdot 1c_2 + 3\cdot 2c_3t + 4\cdot 3c_4t^2 + \cdots)\\ &+ (a_0 + a_1t + a_2t^2 + \cdots)(c_1 + 2c_2t + 3c_3t^2 + \cdots)\\ &+ (b_0 + b_1t + b_2t^2 + \cdots)(c_0 + c_1t + c_2t^2 + \cdots) = 0.\end{aligned}$$

If the series for $y(t)$ converges and if the left side is expanded out, then the coefficients of each power of t must be 0. Thus

$$\begin{aligned}&2\cdot 1c_2 + a_0c_1 + b_0c_0 = 0,\\ &3\cdot 2c_3 + (a_02c_2 + a_1c_1) + (b_0c_1 + b_1c_0) = 0,\\ &4\cdot 3c_4 + (a_03c_3 + a_12c_2 + a_2c_1) + (b_0c_2 + b_1c_1 + b_2c_0) = 0,\\ &\quad\vdots\\ &n(n-1)c_n + (a_0(n-1)c_{n-1} + a_1(n-2)c_{n-2} + \cdots + a_{n-2}c_1)\\ &\qquad\qquad + (b_0c_{n-2} + b_1c_{n-3} + \cdots + b_{n-2}c_0) = 0.\end{aligned}$$

These equations tell us that c_0 and c_1 are arbitrary and that $c_2, c_3, \ldots$ are each determined by the previous coefficients. Thus $c_2, c_3, \ldots$ may be computed inductively. Since $c_0 = y(0)$ and $c_1 = y'(0)$, this degree of flexibility is consistent with the existence and uniqueness theorems.

Theorem 4.14. If $P(t)$ and $Q(t)$ are given by convergent power series for $|t| < R$, then any formal power series that satisfies $y'' + P(t)y' + Q(t)y = 0$ converges for $|t| < R$ to a solution. Consequently every solution of this equation on the interval $-R < t < R$ is given by a power series convergent for $|t| < R$.

PROOF. Fix r with $0 < r < R$, and choose some R_1 with $r < R_1 < R$. Let the notation for the power series of P, Q, and y be as above. Theorem 1.37 shows that the series with terms $|a_nR_1^n|$ and $|b_nR_1^n|$ are convergent, and hence the terms are bounded as functions of n. Thus there exists a real number C such that $|a_n| \le C/R_1^n$ and $|b_n| \le C/R_1^n$ for all $n \ge 0$. We shall show that $|c_n| \le M/r^n$ for a suitable M and all $n \ge 0$.

The constant M will be fixed so that a large initial number of terms have $|c_n| \le M/r^n$, and then we shall see that all subsequent terms satisfy the same inequality. To find an M that works, we start from the formula computed above for c_n:

$$\begin{aligned}n(n-1)c_n = &-(a_0(n-1)c_{n-1} + a_1(n-2)c_{n-2} + \cdots + a_{n-2}c_1)\\ &- (b_0c_{n-2} + b_1c_{n-3} + \cdots + b_{n-2}c_0).\end{aligned}$$

If M works for $0, 1, \ldots, n-1$, then

$$\begin{aligned}
n(n-1)|c_n| &\le CM(n-1)(R_1^{-0}r^{-(n-1)} + R_1^{-1}r^{-(n-2)} + \cdots + R_1^{-(n-2)}r^{-1}) \\
&\quad + CM(R_1^{-0}r^{-(n-2)} + R_1^{-1}r^{-(n-3)} + \cdots + R_1^{-(n-2)}r^{-0}) \\
&= CM(n-1)r^{-n}r\Big(1 + \frac{r}{R_1} + \cdots + \Big(\frac{r}{R_1}\Big)^{n-2}\Big) \\
&\quad + CMr^{-n}r^2\Big(1 + \frac{r}{R_1} + \cdots + \Big(\frac{r}{R_1}\Big)^{n-2}\Big) \\
&\le r^{-n}(CM)(r(n-1) + r^2)\,\frac{1}{1-(r/R_1)}
\end{aligned}$$

and therefore

$$|c_n| \le Mr^{-n}\Big(\frac{CR_1}{R_1 - r}\,\frac{r(n-1)+r^2}{n(n-1)}\Big).$$

For n sufficiently large, the factor in parentheses is ≤ 1. At that point we obtain $|c_n| \le Mr^{-n}$ if $|c_k| \le Mr^{-k}$ for $k < n$, and induction yields the asserted estimate. Thus $\sum c_n t^n$ converges for $|t| < r$. Since r can be arbitrarily close to R, $\sum c_n t^n$ converges for $|t| < R$.

Finally we saw above that c_0 and c_1 are arbitrary and can therefore be matched to any initial data for $y(0)$ and $y'(0)$. Consequently the vector space of power-series solutions convergent for $|t| < R$ has dimension 2. By Theorem 4.6, all solutions on the interval $-R < t < R$ are accounted for. This completes the proof.

As a practical matter, the recursive expression for c_n becomes increasingly complicated as n increases, and a closed-form expression need not be available. However, in certain cases, something special happens that yields a closed-form expression for c_n. Here is an example.

EXAMPLE. **Legendre's equation** is

$$(1-t^2)y'' - 2ty' + p(p+1)y = 0$$

with p a complex constant. To apply the theorem literally, we should first divide the equation by $(1-t^2)$, and then the power-series expansions of the coefficients will be convergent for $|t| < 1$. The theorem says that we obtain two linearly independent power-series solutions of the equation for $|t| < 1$. To compute them, it is more convenient to work with the equation without making the preliminary division. Then the equation gives us

$$\begin{aligned}
&(2c_2 + 3\cdot 2c_3t + 4\cdot 3c_4t^2 + \cdots) - (2c_2t^2 + 3\cdot 2c_3t^3 + 4\cdot 3c_4t^4 + \cdots) \\
&- 2(c_1t + 2c_2t^2 + 3c_3t^3 + 4c_4t^4 + \cdots) + p(p+1)(c_0 + c_1t + c_2t^2 + \cdots) = 0,
\end{aligned}$$

which yields the following formulas for the coefficients:

$$\begin{aligned}
&2c_2 + p(p+1)c_0 = 0,\\
&3 \cdot 2c_3 - 2c_1 + p(p+1)c_1 = 0,\\
&4 \cdot 3c_4 - 2 \cdot 1c_2 - 2 \cdot 2c_2 + p(p+1)c_2 = 0,\\
&\quad\vdots\\
&n(n-1)c_n - [(n-2)(n-3) + 2(n-2) - p(p+1)]c_{n-2} = 0.
\end{aligned}$$

Thus we can write c_n explicitly as a product. We can verify convergence of $\sum c_n t^n$ directly by the ratio test: since

$$\frac{c_n t^n}{c_{n-2} t^{n-2}} = \frac{(n-2)(n-3) + 2(n-2) - p(p+1)}{n(n-1)}\, t^2,$$

we have convergence for $|t| < 1$. Observe that the numerator in the fraction on the right is equal to

$$(n-2)(n-3) + 2(n-2) - p(p+1) = (n-2)(n-1) - p(p+1),$$

and this is 0 when p is an integer ≥ 0 and $n - 2 = p$. Therefore one of the solutions is a polynomial of degree p if p is an integer ≥ 0. Such polynomials, when suitably normalized, are called **Legendre polynomials**.

The second kind of ordinary differential equation for which we shall seek series solutions is

$$t^2 y' + tP(t)y' + Q(t)y = 0,$$

where $P(t)$ and $Q(t)$ are given by convergent power-series expansions for $|t| < R$:

$$\begin{aligned}
P(t) &= a_0 + a_1 t + a_2 t^2 + \cdots,\\
Q(t) &= b_0 + b_1 t + b_2 t^2 + \cdots.
\end{aligned}$$

The existence and uniqueness theorems do not apply to this equation on an interval containing $t = 0$ unless t happens to divide $P(t)$ and t^2 happens to divide $Q(t)$. When this divisibility does not occur, the above equation is said to have a **regular singular point** at $t = 0$. The treatment of the corresponding n^{th}-order equation is no different, but we stick to the second-order case because of its relative importance in applications. For this kind of equation, the treatment of first-order systems is more complicated than the treatment of a single equation of n^{th} order.

Actually, the second-order equation above need not have power series solutions. The prototype for the above equation is the equation

$$t^2y'' + tPy' + Qy = 0$$

with P and Q constant. This equation is known as **Euler's equation** and can be solved in terms of elementary functions. In fact, we make a change of variables by putting $t = e^x$ and $x = \log t$ for $t > 0$. Then we obtain

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} = \frac{1}{t}\frac{dy}{dx}$$

and

$$\frac{d^2y}{dt^2} = -\frac{1}{t^2}\frac{dy}{dx} + \frac{1}{t}\frac{d}{dt}\Big(\frac{dy}{dx}\Big) = -\frac{1}{t^2}\frac{dy}{dx} + \frac{1}{t}\frac{d^2y}{dx^2}\frac{dx}{dt} = -\frac{1}{t^2}\frac{dy}{dx} + \frac{1}{t^2}\frac{d^2y}{dx^2},$$

and hence the equation becomes

$$\frac{d^2y}{dx^2} + (P-1)\,\frac{dy}{dx} + Qy = 0.$$

This is an equation of the kind considered in Section 6. A solution is e^{st}, where s is a root of the characteristic polynomial $s^2 + (P-1)s + Q = 0$. If the two roots of the characteristic polynomial are distinct, we obtain two linearly independent solutions for $x \in (-\infty, +\infty)$, and these transform back to two solutions t^s of the Euler equation for $t > 0$. If the characteristic equation has one root s of multiplicity 2, then we obtain the two linearly independent solutions e^{sx} and xe^{sx} for $x \in (-\infty, +\infty)$, and these transform back to two solutions x^s and $x^s \log x$ for $x > 0$.

In practice, the technique to solve the Euler equation $t^2y'' + tPy' + Qy = 0$ is to substitute $y(t) = t^s$ and obtain $s(s-1)t^s + sPt^s + Qt^s = 0$. This equation holds if and only if s satisfies

$$s(s-1) + sP + Q = 0,$$

which is called the **indicial equation**.

In the general case of a regular singular point, we proceed by analogy and are led to seek for $t > 0$ a series solution of the form

$$y(t) = t^s(c_0 + c_1t + c_2t^2 + \cdots) \qquad \text{with } c_0 \neq 0.$$

Suppose that the power-series part $\sum c_nt^n$ is convergent. We substitute and obtain

$$\begin{aligned} &t^s(c_0s(s-1) + c_1(s+1)st + c_2(s+2)(s+1)t^2 + \cdots) \\ &\quad + t^s(a_0 + a_1t + a_2t^2 + \cdots)(sc_0 + (s+1)c_1t + (s+2)c_2t^2 + \cdots) \\ &\quad + t^s(b_0 + b_1t + b_2t^2 + \cdots)(c_0 + c_1t + c_2t^2 + \cdots) = 0. \end{aligned}$$

Dividing by t^s and setting the coefficient of each power of t equal to 0 gives the equations

$$
\begin{aligned}
&c_0 s(s-1) + sc_0 a_0 + c_0 b_0 = 0,\\
&c_1(s+1)s + ((s+1)c_1 a_0 + sc_0 a_1) + (c_1 b_0 + c_0 b_1) = 0,\\
&c_2(s+2)(s+1) + ((s+2)c_2 a_0 + \cdots) + (c_2 b_0 + \cdots) = 0,\\
&\quad\vdots\\
&c_n(s+n)(s+n-1) + ((s+n)c_n a_0 + \cdots) + (c_n b_0 + \cdots) = 0.
\end{aligned}
$$

Since c_0 is by assumption nonzero, we can divide the first equation by it, and we obtain

$$s(s-1) + a_0 s + b_0 = 0,$$

which is the **indicial equation** for $t^2 y' + tP(t)y' + Q(t)y = 0$. This determines the exponent s. Then c_0 is arbitrary, and all subsequent c_n's can be found recursively, provided the coefficient of c_n in the $(n+1)^{\text{st}}$ equation above is never 0 for $n \geq 1$, i.e., provided

$$(s+n)(s+n+1) + (s+n)a_0 + b_0 \neq 0 \qquad \text{for } n \geq 1.$$

In other words, we can solve recursively for all c_n in terms of c_0 provided $s+n$ does not satisfy the indicial equation for any $n \geq 1$. We summarize as follows.

Proposition 4.15. If $P(t)$ and $Q(t)$ are given by convergent power series for $|t| < R$, then the following can be said about formal series solutions of $t^2 y'' + tP(t)y' + Q(t)y = 0$ of the type $t^s(c_0 + c_1 t + c_2 t^2 + \cdots)$ with $c_0 \neq 0$:

(a) If the indicial equation has distinct roots not differing by an integer, then there are formal solutions of the type $x^s(c_0 + c_1 t + c_2 t^2 + \cdots)$ for each root s of the indicial equation.

(b) If the indicial equation has roots $r_1 \leq r_2$ with $r_2 - r_1$ equal to an integer, then there is a 1-parameter family of formal solutions of the type $t^{r_2}(c_0 + c_1 t + c_2 t^2 + \cdots)$ with $c_0 \neq 0$. If $r_1 < r_2$ in addition, there may be formal solutions $t^{r_1}(c_0 + c_1 t + c_2 t^2 + \cdots)$ with $c_0 \neq 0$, as there are for an Euler equation.

Theorem 4.16. If $P(t)$ and $Q(t)$ are given by convergent power series for $|t| < R$, then all formal series solutions of $t^2 y'' + tP(t)y' + Q(t)y = 0$ of the type $t^s(c_0 + c_1 t + c_2 t^2 + \cdots)$ with $c_0 \neq 0$ converge for $0 < t < R$ to a function that is a solution for $0 < t < R$.

PROOF. As in the proof of Theorem 4.14, fix r with $0 < r < R$, and choose some R_1 with $r < R_1 < R$. Let the series expansions of $P(t)$ and $Q(t)$ be as above, so that there is a number C with $|a_n| \le C/R_1^n$ and $|b_n| \le C/R_1^n$. Choose N large enough so that

$$\frac{Cr/R_1}{1-r/R_1}\left(\frac{|s|+n+1}{|(s+n)(s+n+1)+a_0(s+n)+b_0|}\right) \le 1 \qquad (*)$$

for $n \ge N$. Then choose M such that $|c_n| \le M/r^n$ for $n \le N$. We shall prove by induction on n that $|c_n| \le M/r^n$ for all n. The base case of the induction is $n = N$, where the inequality holds by definition of M. Suppose it holds for $1, \dots, n-1$. The formula for c_n is

$$\begin{aligned} &c_n\big((s+n)(s+n-1)+a_0(s+n)+b_0\big) \\ &\qquad\qquad = -[(s+n-1)a_1c_{n-1}+\cdots+sa_nc_0]-[b_1c_{n-1}+\cdots+b_nc_0]. \end{aligned}$$

Our inductive hypothesis gives

$$\begin{aligned} |c_n||(s+n)(s+n-1)+a_0(s+n)+b_0| &\\ \le CM(|s|+n)&(R_1^{-1}r^{-(n-1)}+\cdots+R_1^{-n}r^0) \\ &+CM(R_1^{-1}r^{-(n-1)}+\cdots+R_n^{-n}r^0) \\ = CM(|s|+n&+1)r^{-n}\Big(\frac{r}{R_1}+\cdots+\frac{r^n}{R_1^n}\Big) \\ \le Mr^{-n}&\left[C(|s|+n+1)\left(\frac{r/R_1}{1-r/R_1}\right)\right]. \end{aligned}$$

Thus

$$|c_n| \le Mr^{-n}\left[\frac{Cr/R_1}{1-r/R_1}\left(\frac{|s|+n+1}{|(s+n)(s+n+1)+a_0(s+n)+b_0|}\right)\right] \le Mr^{-n},$$

the second inequality holding by $(*)$, and the induction is complete.

It follows that $\sum c_nt^n$ converges for $|t| < r$. Since r can be arbitrarily close to R, $\sum c_nt^n$ converges for $|t| < R$. This completes the proof.

EXAMPLE. **Bessel's equation** of order p with $p \ge 0$. This is the equation

$$t^2y'' + ty' + (t^2 - p^2)y = 0.$$

It has $P(t) = 1$ and $Q(t) = t^2 - p^2$, both with infinite radius of convergence. The indicial equation in general is $s(s-1) + a_0s + b_0 = 0$ and hence is

$$s(s-1) + s - p^2 = 0$$

in this case. Thus $s = \pm p$. Theorem 4.16 shows that there is a solution of the form

$$J_p(t) = t^p\Big(\frac{1}{2^p p!} + c_1 t + c_2 t^2 + \cdots\Big),$$

and this is defined to be the **Bessel function** of order p. The theorem gives another solution of the form t^{-p} times a power series except possibly when p is an integer or a half integer. To determine all these solutions, we substitute the series $t^s \sum c_n t^n$ and get

$$\begin{aligned}
&s(s-1)c_0 + (s+1)sc_1 t + (s+2)(s+1)c_2 t^2 + \cdots \\
&+ \quad sc_0 + (s+1)c_1 t + (s+2)c_2 t^2 + \cdots \\
&+ \quad c_0 t^2 + c_1 t^3 + \cdots \\
&- \quad p^2 c_0 - p^2 c_1 t - p^2 c_2 t^2 - p^2 c_3 t^3 - \cdots = 0.
\end{aligned}$$

The resulting equations are

$$\begin{aligned}
&[s(s-1) + s - p^2]c_0 = 0 && \text{from } t^0, \\
&[(s+1)s + (s+1) - p^2]c_1 = 0 && \text{from } t^1, \\
&[(s+n)(s+n-1) + (s+n) - p^2]c_n + c_{n-2} = 0 && \text{from } t^n \text{ for } n \geq 2.
\end{aligned}$$

The first of these equations repeats the indicial equation, giving $s = \pm p$. The second says that either $c_1 = 0$ or that $s+1$ solves the indicial equation. In the latter case $s = -\frac{1}{2}$ and $p = \frac{1}{2}$. The third says that $[(s+n)^2 - p^2]c_n = -c_{n-2}$. For the case that $s = +p$, we obtain

$$c_n = \frac{-c_{n-2}}{(p+n)^2 - p^2},$$

and there is no problem from the denominator. The result is that the Bessel function of order $p \geq 0$ is given by

$$J_p(t) = \frac{t^p}{2^p p!}\Big(1 - \frac{t^2}{2(2p+2)} + \frac{t^4}{2\cdot 4(2p+2)(2p+4)} - \cdots\Big).$$

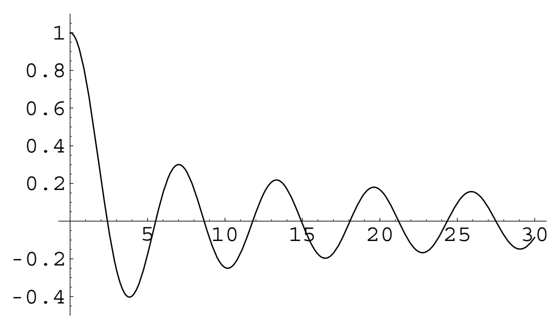


FIGURE 4.3. Graph of Bessel function $J_0(t)$.

For the case that $s = -p$, we obtain

$$c_n = \frac{-c_{n-2}}{(-p+n)^2 - p^2},$$

and the denominator gives a problem for $n = 2p$ and for no other value of n. If p is an integer, the problematic n is even and we must have $c_{n-2} = 0, c_{n-4} = 0, \dots, c_0 = 0$. The condition $c_0 = 0$ is a contradiction, and we conclude that there is no solution of the form t^{-p} times a nonzero power series; indeed, Problems 18–19 at the end of the chapter will identify a different kind of solution. If p is a half integer but not an integer, then the problematic n is odd, and we are led to conclude that $0 = c_{n-2} = \cdots = c_3 = c_1$, with c_0 and c_{2p} arbitrary. There is no contradiction, and we obtain a solution of the form t^{-p} times a nonzero power series.

9. Problems

1. For the differential equation $yy' = -t$:
   (a) Solve the equation.
   (b) Find all points (t_0, y_0) where the the existence theorem and the uniqueness theorem of Section 2 do not apply.
   (c) For each point (t_0, y_0) not in (b), give a solution $y(t)$ with $y(t_0) = y_0$.

2. Prove that the equation $y' = t + y^2$ has a solution satisfying the initial condition $y(0) = 0$ and defined for $|t| < 1/2$.

3. In classical notation, a particular vector field in the plane is given by $\sqrt{x}\,\frac{\partial}{\partial x} + \frac{1}{2}\frac{\partial}{\partial y}$. Find a parametric realization of an integral curve for this vector field passing through $(1, 1)$.

4. Evaluate $\displaystyle \frac{d}{dt}\int_0^{t^2} \frac{1}{s}(\sin st)\,ds$.

5. Find all solutions on $(-\infty, +\infty)$ to $y'' - 3y' + 2y = 4$.

6. (a) For each of these matrices A, find matrices B and J, with J in Jordan form, such that $A = BJB^{-1}$: $\quad A = \begin{pmatrix} 1 & 1 \\ 4 & -5 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$
   (b) For each of the matrices A in (a), find a basis of solutions $y(t)$ to the system of differential equations $y' = Ay$.

7. The n^{th}-order equation $y^{(n)}+a_{n-1}y^{(n-1)}+\cdots+a_0y = 0$ with constant coefficients leads to a linear system $z' = Az$ with

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ & 0 & 1 & 0 & \cdots & 0 & 0 \\ & & 0 & 1 & \cdots & 0 & 0 \\ & & & \ddots & \ddots & \vdots & \vdots \\ & & & & 0 & 1 & 0 \\ & & & & & 0 & 1 \\ -a_0 & -a_1 & -a_2 & & \cdots & & -a_{n-1} \end{pmatrix}.$$

Prove that $\det(\lambda 1 - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0$ by expanding the determinant by cofactors.

8. (a) Let $\{f_n\}$ be a uniformly bounded sequence of Riemann integrable functions on $[0, 1]$. Define $F_n(t) = \int_0^t f_n(s)\,ds$. Prove that $\{F_n\}$ is an equicontinuous family of functions on $[0, 1]$.
   (b) Prove that the set of functions $y(t)$ on $[0, 1]$ with $y'' + y = f(t)$ and $y(0) = y'(0) = 0$ is equicontinuous as f varies over the set of continuous functions on $[0, 1]$ with $0 \le f(t) \le 1$ for all t.
   (c) Let $u(t)$ be continuous on $[a, b]$. Prove that the set of functions $y(t)$ on $[a, b]$ with $y'' + q(t)y = f(t)$ and $y(0) = y'(0) = 0$ is equicontinuous as $f(t)$ varies over the set of continuous functions on $[0, 1]$ with $0 \le f(t) \le 1$ for all t.

9. The differential equation $t^2y'' + (3t-1)y' + y = 0$ has an irregular singular point at $t = 0$.
   (a) Verify that $\sum_{n=0}^{\infty}(n!)t^n$ is a formal power series solution of the equation even though the power series has radius of convergence 0.
   (b) Verify that $y(t) = t^{-1}e^{-1/t}$ is a solution for $t > 0$.

Problems 10–13 concern harmonic functions in the open unit disk, which were introduced in Problems 14–15 at the end of Chapter III. The first objective here is to use ordinary differential equations and Fourier series to show that all these functions may be expressed in a relatively simple form. The second objective is to use convolution, as defined in Problem 8 at the end of Chapter III, to relate this formula to the Poisson kernel, which was defined in Problems 27–29 at the end of Chapter I. Problems 10–12 here are an instance of the method of **separation of variables**, a beginning technique with partial differential equations; this topic is developed further in the companion volume *Advanced Real Analysis*. In all problems in this set, let $u(x, y)$ be harmonic in the open unit disk.

10. Write $u(x, y)$ in polar coordinates as $u(r\cos\theta, r\sin\theta) = v(r, \theta)$. Using Fourier series, show for $0 \le r < 1$ and any $\delta > 0$ that $v(r, \theta)$ is the sum of an absolutely convergent Fourier series $\sum_{n=-\infty}^{\infty} c_n(r)e^{in\theta}$ with $|c_n(r)| \le M/n^2$ for $0 \le r \le 1 - \delta$ for some M depending on δ.

11. Let R_θ be the rotation matrix defined in Problem 15 at the end of Chapter III. That problem shows that $(u \circ R_\varphi)(x, y) = v(r, \theta + \varphi)$ is harmonic for each φ. Prove that $\frac{1}{2\pi}\int_{-\pi}^{\pi}(u \circ R_\varphi)(x, y)e^{-ik\varphi}\,d\varphi$ is harmonic and is given in polar coordinates by $c_k(r)e^{ik\theta}$.
12. By computing with the Laplacian in polar coordinates and showing that $c_k(r)$ is bounded as $r \downarrow 0$, prove that $c_k(r) = a_k r^{|k|}$ for some complex constant a_k. Conclude that every harmonic function in the open unit disk is of the form $v(r, \theta) = \sum_{n=-\infty}^{\infty} c_n r^{|n|} e^{in\theta}$, the sum being absolutely convergent for all r with $0 \le r < 1$.
13. Deduce from Problem 8 at the end of Chapter III that if $v(r, \theta)$ is as in the previous problem and if $0 < R < 1$, then $v(r, \theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_R(\varphi)P_{r/R}(\theta - \varphi)\,d\varphi$ for $0 \le r < R$, where P is the Poisson kernel and f_R is the C^∞ function $f_R(\theta) = \sum_{n=-\infty}^{\infty} c_n R^{|n|} e^{in\theta}$.

Problems 14–17 concern homogeneous linear differential equations. Except for the first of the problems, each works with a substitution in a second-order equation that simplifies the equation in some way.

14. If $a(t)$ is continuous on an interval and $A(t)$ is an indefinite integral, verify that all solutions of the single first-order linear homogeneous equation $y' = a(t)y$ are of the form $y(t) = ce^{A(t)}$.
15. (a) Suppose that $u(t)$ is a nowhere vanishing solution of

$$y'' + P(t)y' + Q(t)y = 0$$

on an interval, with P and Q assumed continuous. Look for a solution of the form $u(t)v(t)$, and derive the necessary and sufficient condition

$$v'(t) = cu(t)^{-2}e^{-\int P(t)\,dt}.$$

(b) For $y'' - ty' - y = 0$, one solution is $e^{t^2/2}$. Find a linearly independent solution.
16. Let $y'' + P(t)y' + Q(t)y = 0$ be given with P, P', and Q continuous on an interval. Write $y(t) = u(t)v(t)$, substitute, regard $u(t)$ as known, and obtain a second-order equation for v. Show how to choose $u(t)$ to make the coefficient of v' be 0, and thus reduce the given equation to an equation $v'' + R(t)v = 0$ with R continuous. Give a formula for R.
17. If $L(v) = (pv')' - qv + \lambda rv$, show that the substitution $u = v\sqrt{r}$ changes $L(v) = 0$ into $L_0(u) = 0$, where $L_0(u) = (p^*u')' - q^*u + \lambda u$ with $p^* = p/r$.

Problems 18–19 concern finding the form of the second solution to a second-order equation with a regular singular point. The first of the two problems amounts to a result in complex analysis but requires nothing beyond Chapter I of this book.

18. Suppose that $\sum_{n=0}^{\infty} c_n x^n$ is a power series with $c_0 = 1$.
    (a) Write down recursive formulas for the coefficients d_n of a power series $\sum_{n=0}^{\infty} d_n x^n$ with $d_0 = 1$ such that $\big(\sum_{n=0}^{\infty} c_n x^n\big)\big(\sum_{n=0}^{\infty} d_n x^n\big) = 1$.
    (b) Prove, by induction on n, that if $|c_n| \le Mr^n$ for all $n \ge 0$, then $|d_n| \le M(M+1)^{n-1} r^n$ for all $n \ge 1$.
    (c) Prove that if $f(0) \ne 0$ and if $f(x)$ is the sum of a convergent power series for $|x| < R$ for some $R > 0$, then $1/f(x)$ is the sum of a convergent power series for $|x| < \varepsilon$ for some $\varepsilon > 0$.

19. Suppose that $P(t)$ and $Q(t)$ are given near $t = 0$ by power series with positive radii of convergence. Take for granted that if $a(t)$ is given by a power series with a positive radius of convergence, then so is $e^{a(t)}$. Form the equation

$$t^2 y'' + tP(t)y' + Q(t)y = 0,$$

let s_1 and s_2 be the two roots of the indicial equation, and suppose that the differential equation has a solution given on some interval $(0, \varepsilon)$ by $f(t) = t^{s_1} \sum_{n=0}^{\infty} c_n t^n$ with $c_0 \ne 0$.
    (a) Using Problem 15a, prove that the differential equation has a linearly independent solution given on some interval $(0, \varepsilon')$ by

$$g(t) = cf(t)\log t + t^{s_2} \sum_{n=0}^{\infty} k_n t^n \qquad \text{with } k_0 \ne 0.$$

    (b) Prove that the coefficient c in $g(t)$ is $\ne 0$ if $s_1 = s_2$.
    (c) For Bessel's equation $t^2 y'' + ty' + (t^2 - p^2)y = 0$ with $p \ge 0$ an integer and with $s_1 = p$ and $s_2 = -p$, show that the coefficient c in $g(t)$ is $\ne 0$. Thus there is a solution of the form $J_p(t)\log t + t^{-p}$(power series) on some interval $(0, \varepsilon')$.

Problems 20–25 prove the **Cauchy–Peano Existence Theorem**, that a local solution in Theorem 4.1 to $y' = F(t, y)$ and $y(t_0) = y_0$ exists if F is continuous even if F does not satisfy a Lipschitz condition. The idea is to construct a sequence of polygonal approximations to solutions, check that they form an equicontinuous family, apply Ascoli's Theorem (Theorem 2.56) to extract a uniformly convergent subsequence, and then see that the limit of the subsequence is a solution. A member of the sequence of polygonal approximations depends on a number $\epsilon > 0$. With notation as in the statement of Theorem 4.1, the construction for $[t_0, t_0 + a']$ is as follows: Choose the δ of uniform continuity for F and ϵ on the set R. Fix a partition $t_0 < t_1 < \cdots < t_n = t_0 + a'$ of $[t_0, t_0 + a']$ with $\max_k\{t_k - t_{k-1}\} \le \min(\delta, \delta/M)$. Define $y(t)$, as a function of ϵ, for $t_{k-1} < t \le t_k$ inductively on k by $y(t_0) = y_0$ and

$$y(t) = y(t_{k-1}) + F(t_{k-1}, y(t_{k-1}))(t - t_{k-1}).$$

20. Check that the formula for $y(t)$ when $t_{k-1} < t \le t_k$ remains valid when $t = t_{k-1}$, and conclude that $y(t)$ is continuous. Then prove by induction on k that $|y(t) - y(t_0)| \le M(t - t_0) \le b$ for $t_{k-1} \le t \le t_k$, and deduce that $(t, y(t))$ is in R' for $t_0 \le t \le t_0 + a'$.
21. Prove that $|y(t) - y(t')| \le M|t - t'|$ if t and t' are both in $[t_0, t_0 + a']$.
22. The function $y'(t)$ is defined on $[t_0, t_0 + a']$ except at the points of the partition and is given by $y'(t) = F(t_{k-1}, y(t_{k-1}))$ if $t_{k-1} < t < t_k$. Prove that $y(t) = y_0 + \int_{t_0}^t y'(s)\,ds$ for $t_0 \le t \le t_0 + a'$ and that $|y'(s) - F(s, y(s))| \le \epsilon$ if $t_{k-1} < s < t_k$.
23. Writing $y(t) = y_0 + \int_{t_0}^t [F(s, y(s)) + [y'(s) - F(s, y(s))]]\,ds$ and using the result of the previous problem, prove for all t in $[t_0, t_0 + a']$ that
$$\Big| y(t) - \big(y_0 + \textstyle\int_{t_0}^t F(s, y(s))\,ds\big)\Big| \le \epsilon a'.$$
24. Let ϵ_n be a monotone decreasing sequence with limit 0, and let $y_n(t)$ be a function for t in $[t_0, t_0 + a']$ constructed as above for the number ϵ_n. Deduce from Problem 21 that $\{y_n(t)\}$ is uniformly bounded and uniformly equicontinuous for t in $[t_0, t_0 + a']$.
25. Apply Ascoli's Theorem to $\{y_n\}$, and let $y(t)$ be the uniform limit of a uniformly convergent subsequence of $\{y_n\}$. Prove that $y(t)$ is continuous, and use Problem 23 to prove that $y(t) = y_0 + \int_{t_0}^t F(s, y(s))\,ds$. What modifications are needed to the argument to handle $[t_0 - a', t_0]$?

CHAPTER V

Lebesgue Measure and Abstract Measure Theory

Abstract. This chapter develops the basic theory of measure and integration, including Lebesgue measure and Lebesgue integration for the line.

Section 1 introduces measures, including 1-dimensional Lebesgue measure as the primary example, and develops simple properties of them. Sections 2–4 introduce measurable functions and the Lebesgue integral and go on to establish some easy properties of integration and the fundamental theorems about how Lebesgue integration behaves under limit operations.

Sections 5–6 concern the Extension Theorem announced in Section 1 and used as the final step in the construction of Lebesgue measure. The theorem allows σ-finite measures to be extended from algebras of sets to σ-algebras. The theorem is proved in Section 5, and the completion of a measure space is defined in Section 6 and related to the proof of the Extension Theorem.

Section 7 treats Fubini's Theorem, which allows interchange of order of integration under rather general circumstances. This is a deep result. As part of the proof, product measure is constructed and important measurability conditions are established. This section mentions that Fubini's Theorem will be applicable to higher-dimensional Lebesgue measure, but the details are deferred to Chapter VI.

Section 8 extends Lebesgue integration to complex-valued functions and to functions with values in finite-dimensional vector spaces.

Section 9 gives a careful definition of the spaces L^1, L^2, and L^∞ for any measure space, introduces the notion of a normed linear space, and verifies that these three spaces are examples. The main theorem of the section about L^1, L^2, and L^∞ is the completeness of these three spaces as metric spaces. In addition, the section proves a version of Alaoglu's Theorem concerning weak-star convergence.

1. Measures and Examples

In the theory of the Riemann integral, as discussed in Chapter I for $\mathbb{R}^1$ and in Chapter III for $\mathbb{R}^n$, we saw that Riemann integration is a powerful tool when applied to continuous functions. Riemann integration makes sense also when applied to certain kinds of discontinuous functions, but then the theory has some weaknesses.

Without any change in the definitions, one of these is that the theory applies only to bounded functions. Thus we can compute $\int_0^1 x^p\,dx = \big[x^{p+1}/(p+1)\big]_0^1 = (p+1)^{-1}$ for $p \geq 0$, but only the right side makes sense for $-1 < p < 0$. More seriously we made calculations with trigonometric series in Section I.10 and found that $\frac{1}{2}\log\big(\frac{1}{2-2\cos\theta}\big) = \sum_{n=1}^\infty \frac{\cos n\theta}{n}$ and $\frac{1}{2}(\pi-\theta) = \sum_{n=1}^\infty \frac{\sin n\theta}{n}$ for $0 < \theta < 2\pi$.

When we tried to explain these similar-looking identities with Fourier series, we were able to handle the second one because $\frac{1}{2}(\pi - \theta)$ is a bounded function, but we were not able to handle the first one because $\frac{1}{2}\log\left(\frac{1}{2-2\cos\theta}\right)$ is unbounded.

Other weaknesses appeared in Chapters I–IV at certain times: when we always had to arrange for the set of integration to be bounded, when we had no clue which sequences $\{c_n\}$ of Fourier coefficients occurred in the beautiful formula given by Parseval's Theorem, when Fubini's Theorem turned out to be awkward to apply to discontinuous functions, and when the change-of-variables formula did not immediately yield the desired identities even in simple cases like the change from Cartesian coordinates to polar coordinates.

The Lebesgue integral will solve all these difficulties when formed with respect to "Lebesgue measure" in the setting of $\mathbb{R}^n$. In addition, the Lebesgue integral will be meaningful in other settings. For example, the Lebesgue integral will be meaningful on the unit sphere in Euclidean space, while the Riemann integral would always require a choice of coordinates. The Lebesgue integral will be meaningful also in other situations where we can take advantage of some action by a group (such as a rotation group) that is difficult to handle when the setting has to be Euclidean. And the Lebesgue integral will enable us to provide a rigorous foundation for the theory of probability.

There are five ingredients in Lebesgue integration, and these will be introduced in Sections 1–3 of this chapter:

(i) an underlying nonempty set, such as $\mathbb{R}^1$ in the case of Lebesgue integration on the line,
(ii) a distinguished class of subsets, called the "measurable sets," which will form a "σ-ring" or a "σ-algebra,"
(iii) a measure, which attaches a member of $[0, +\infty]$ to each measurable set and which will be "length" in the case of Lebesgue measure on the line,
(iv) the "measurable functions," those functions with values in $\mathbb{R}$ (or some more general space) that we try to integrate,
(v) the integral of a measurable function over a measurable set.

Let us write X for the underlying nonempty set. The important thing about whatever sets are measurable will be that certain simple set-theoretic operations lead from measurable sets to measurable sets. The two main definitions are those of an "algebra" of sets and a "σ-algebra," but we shall refer also to the notions of a "ring" of sets and a "σ-ring" in order to simplify certain technical problems in constructing measures. An **algebra of sets** $\mathcal{A}$ is a set of subsets of X containing $\varnothing$ and X and closed under the operation of forming the union $E \cup F$ of two sets and under taking the complement E^c of a set. An algebra is necessarily closed under intersection $E \cap F$ and difference $E - F = E \cap F^c$. Another operation under which $\mathcal{A}$ is closed is **symmetric difference**, which is defined by

$E \,\Delta\, F = (E - F) \cup (F - E)$; we shall make extensive use of this operation[1] in Section 6 of this chapter.

In practice, despite the effort often needed to define an interesting measure on the sets in an algebra, the closure properties[2] of the algebra are insufficient to deal with questions about limits. For this reason one defines a **σ-algebra** of subsets of X to be an algebra that is closed under countable unions (and hence also countable intersections). Typically a general foundational theorem (Theorem 5.5 below) is used to extend the constructed would-be measure from an algebra to a σ-algebra.

A **ring** $\mathcal{R}$ of subsets of X is a set of subsets closed under finite unions and under difference. Then $\mathcal{R}$ is closed also under the operations of finite intersections, difference, and symmetric difference.[3] A **σ-ring** of subsets of X is a ring of subsets that is closed under countable unions.

EXAMPLES.

(1) $\mathcal{A} = \{\varnothing, X\}$. This is a σ-algebra.

(2) All subsets of X. This is a σ-algebra.

(3) All finite subsets of X. This is a ring. If the complements of such sets are included, the result is an algebra.

(4) All finite and countably infinite subsets of X. This is a σ-ring. If the complements of such sets are included, the result is a σ-algebra.

(5) All **elementary sets** of $\mathbb{R}$. These are all finite disjoint unions of bounded intervals in $\mathbb{R}$ with or without endpoints. This collection is a ring. To see the closure properties, we first verify that any finite union of bounded intervals is a finite disjoint union; in fact, if $I_1, \ldots, I_n$ are bounded intervals such that none contains any of the others, then $I_k - \bigcup_{m=1}^{k-1} I_m$ is an interval, and these intervals are disjoint as k varies; also these intervals have the same union as $I_1, \ldots, I_n$. Now let $E = \bigcup_i I_i$ and $F = \bigcup_j J_j$ be given. Since $I_i \cap J_j$ is an interval, the identity $E \cap F = \bigcup_{i,j} (I_i \cap J_j)$ shows that $E \cap F$ is a finite union of intervals. Since each $I_i - J_j$ is an interval or the union of two intervals, the identity $E - F = \bigcup_i \bigcap_j (I_i - J_j)$ then shows that $E - F$ is a finite union of intervals.

(6) If $\mathcal{C}$ is an arbitrary class of subsets of X, then there is a unique smallest algebra $\mathcal{A}$ of subsets of X containing $\mathcal{C}$. Similar statements apply to σ-algebras,

---

[1]For some properties of symmetric difference, see Problem 1 at the end of the chapter.

[2]An algebra of sets really is an algebra in the sense of the discussion of algebras with the Stone–Weierstrass Theorem (Theorem 2.58). The scalars replacing $\mathbb{R}$ or $\mathbb{C}$ are the members of the two-element field $\{0, 1\}$, addition is given by symmetric difference, and multiplication is given by intersection. The additive identity is $\varnothing$, the multiplicative identity is X, and every element is its own negative. Multiplication is commutative.

[3]A ring of sets really is a ring in the sense of modern algebra; addition is given by symmetric difference, and multiplication is given by intersection.

rings, and σ-rings. In fact, consider all algebras of subsets of X containing $\mathcal{C}$. Example 2 shows that there is one. Let $\mathcal{A}$ be the intersection of all these algebras, i.e., the set of all subsets that occur in each of these algebras. If two sets occur in $\mathcal{A}$, they occur in each such algebra, and their intersection is in each algebra. Hence their intersection is in $\mathcal{A}$. Similarly $\mathcal{A}$ is closed under differences.

If $\mathcal{R}$ is a ring of subsets of X, a **set function** is a function $\rho : \mathcal{R} \to \mathbb{R}^*$, where $\mathbb{R}^*$ denotes the extended real-number system as in Section I.1. The set function is **nonnegative** if its values are all in $[0, +\infty]$, it is **additive** if $\rho(\varnothing) = 0$ and if $\rho(E \cup F) = \rho(E) + \rho(F)$ whenever E and F are disjoint sets in $\mathcal{R}$, and it is **completely additive** or **countably additive** if $\rho(\varnothing) = 0$ and if $\rho\big(\bigcup_{n=1}^\infty E_n\big) = \sum_{n=1}^\infty \rho(E_n)$ whenever the sets E_n are pairwise disjoint members of $\mathcal{R}$ with $\bigcup_{n=1}^\infty E_n$ in $\mathcal{R}$. In the definitions of "additive" and "completely additive," it is taken as part of the definition that the sums in question are to be well defined in $\mathbb{R}^*$. Observe that completely additive implies additive, since $\rho(\varnothing) = 0$.

Proposition 5.1. An additive set function ρ on a ring $\mathcal{R}$ of sets has the following properties:

(a) $\rho\big(\bigcup_{n-1}^N E_n\big) = \sum_{n=1}^N \rho(E_n)$ if the sets E_n are pairwise disjoint and are in $\mathcal{R}$.

(b) $\rho(E \cup F) + \rho(E \cap F) = \rho(E) + \rho(F)$ if E and F are in $\mathcal{R}$.

(c) If E and F are in $\mathcal{R}$ and $|\rho(E)| < +\infty$, then $|\rho(E \cap F)| < +\infty$.

(d) If E and F are in $\mathcal{R}$ and if $|\rho(E \cap F)| < +\infty$, then $\rho(E - F) = \rho(E) - \rho(E \cap F)$.

(e) If ρ is nonnegative and if E and F are in $\mathcal{R}$ with $E \subseteq F$, then $\rho(E) \le \rho(F)$.

(f) If ρ is nonnegative and if $E, E_1, \dots, E_N$ are sets in $\mathcal{R}$ such that $E \subseteq \bigcup_{n=1}^N E_n$, then $\rho(E) \le \sum_{n=1}^N \rho(E_n)$.

(g) If ρ is nonnegative and completely additive and if $E, E_1, E_2, \dots$ are sets in $\mathcal{R}$ such that $E \subseteq \bigcup_{n=1}^\infty E_n$, then $\rho(E) \le \sum_{n=1}^\infty \rho(E_n)$.

PROOF. Part (a) follows by induction from the definition. In (b), we have $E \cup F = (E - F) \cup (E \cap F) \cup (F - E)$ disjointly. Application of (a) gives $\rho(E \cup F) = \rho(E - F) + \rho(E \cap F) + \rho(F - E)$, with $+\infty$ and $-\infty$ not both occurring. Adding $\rho(E \cap F)$ to both sides, regrouping terms, and taking into account that $\rho(E) = \rho(E-F) + \rho(E \cap F)$ and $\rho(F) = \rho(F-E) + \rho(E \cap F)$, we obtain (b). The right side of the identity $\rho(E) = \rho(E \cap F) + \rho(E - F)$ cannot be well defined if $\rho(E)$ is finite and $\rho(E \cap F)$ is infinite, and thus (c) follows. In the identity $\rho(E) = \rho(E \cap F) + \rho(E - F)$, we can subtract $\rho(E \cap F)$ from both sides and obtain (d) if $\rho(E \cap F)$ is finite. For (e), the inclusion $E \subseteq F$ forces $F = (F - E) \cup E$ disjointly; then $\rho(F) = \rho(F - E) + \rho(E)$, and (e) follows. In (f), put $F_n = E_n - \bigcup_{k=1}^{n-1} E_k$. Then $E = \bigcup_{n=1}^N (E \cap F_n)$ disjointly, and (a) and

(e) give $\rho(E) = \sum_{n=1}^N \rho(E \cap F_n) \le \sum_{n=1}^N \rho(F_n) \le \sum_{n=1}^N \rho(E_n)$. Conclusion (g) is proved in the same way as (f).

Proposition 5.2. Let ρ be an additive set function on a ring $\mathcal{R}$ of sets. If ρ is completely additive, then $\rho(E) = \lim \rho(E_n)$ whenever $\{E_n\}$ is an increasing sequence of members of $\mathcal{R}$ with union E in $\mathcal{R}$. Conversely if $\rho(E) = \lim \rho(E_n)$ for all such sequences, then ρ is completely additive.

PROOF. First we prove the direct part of the proposition. For E and E_n as in the statement, let $F_1 = E_1$ and $F_n = E_n - E_{n-1}$ for $n \ge 2$. Then $E_n = \bigcup_{k=1}^n F_k$ disjointly, and $\rho(E_n) = \sum_{k=1}^n \rho(F_k)$ by additivity. Also, $E = \bigcup_{k=1}^\infty F_k$, and complete additivity gives $\rho(E) = \sum_{k=1}^\infty \rho(F_k) = \lim \sum_{k=1}^n \rho(F_k) = \lim \rho(E_n)$. The direct part of the proposition follows.

For the converse let $\{F_n\}$ be a disjoint sequence in $\mathcal{R}$ with union F in $\mathcal{R}$. Put $E_n = \bigcup_{k=1}^n F_k$. Then E_n is an increasing sequence of sets in $\mathcal{R}$ with union F in $\mathcal{R}$. We are given that $\rho(F) = \lim \rho(E_n)$, and we have $\rho(E_n) = \sum_{k=1}^n \rho(F_k)$ by additivity and Proposition 5.1a. Therefore $\rho(F) = \sum_{k=1}^\infty \rho(F_k)$, and we conclude that ρ is completely additive.

Corollary 5.3. Let ρ be an additive set function on an algebra $\mathcal{A}$ of subsets of X such that $|\rho(X)| < +\infty$. If ρ is completely additive, then $\rho(E) = \lim \rho(E_n)$ whenever $\{E_n\}$ is a decreasing sequence of members of $\mathcal{A}$ with intersection E in $\mathcal{A}$. Conversely if $\lim \rho(E_n) = 0$ whenever $\{E_n\}$ is a decreasing sequence of members of $\mathcal{A}$ with intersection empty, then ρ is completely additive.

PROOF. This follows from Proposition 5.2 by taking complements.

A **measure** is a nonnegative completely additive set function on a σ-ring of subsets of X. If no ambiguity is possible about the σ-ring, we may refer to a "measure on X." When we use measures to work with integrals, the σ-ring will be taken to be a σ-algebra; if integration were to be defined relative to a σ-ring that is not a σ-algebra, then nonzero constant functions would not be measurable.

The assumption that our σ-ring is a σ-algebra for doing integration is no loss of generality. Even when the σ-ring is not a σ-algebra, there is a canonical way of extending a measure from a σ-ring to the smallest σ-algebra containing the σ-ring. Proposition 5.37 at the end of Section 5 gives the details.

EXAMPLES.

(1) For $\{\varnothing, X\}$, define $\mu(X) = a \ge 0$. This is a measure.

(2) For X equal to a countable set and with all subsets in the σ-algebra, attach a weight ≥ 0 to each member of X. Define $\mu(E)$ to be the sum of the weights for the members of E. This is a measure.

(3) For X arbitrary but nonempty, let $\mu(E)$ be the number of points in E, a nonnegative integer or $+\infty$. We refer to μ as **counting measure**.

(4) **Lebesgue measure** m on the ring $\mathcal{R}$ of elementary sets of $\mathbb{R}$. If E is a finite disjoint union of bounded intervals, we let $m(E)$ be the sum of the lengths of the intervals. We need to see that this definition is unambiguous. Consider the special case that $J = I_1 \cup \cdots \cup I_r$ disjointly with I_k extending from a_k to b_k, with or without endpoints. Then we can arrange the intervals in order so that $b_k = a_{k+1}$ for $k = 1, \dots, r-1$. In this case, $m(J) = b_r - a_1$ and $\sum_{k=1}^r m(I_k) = \sum_{k=1}^r (b_k - a_k) = b_r - a_1$. Thus the definition is unambiguous in this special case. If $E = I_1 \cup \cdots \cup I_r = J_1 \cup \cdots \cup J_s$, then the special case gives $m(J_k) = \sum_{j=1}^r m(I_j \cap J_k)$ and hence $\sum_{k=1}^s m(J_k) = \sum_{j,k} m(I_j \cap J_k)$. Reversing the roles of the I_j's and the J_k's, we obtain $\sum_{j=1}^r m(I_j) = \sum_{j,k} m(I_j \cap J_k)$. Thus $\sum_{k=1}^s m(J_k) = \sum_{j=1}^r m(I_j)$, and the definition of m on $\mathcal{R}$ is unambiguous. It is evident that m is nonnegative and additive. We shall prove that m is completely additive on $\mathcal{R}$. Even so, m will not yet be a measure, since $\mathcal{R}$ is not a σ-ring. That step will have to be carried out separately. Proving that m is completely additive on the ring $\mathcal{R}$ uses the fact that m is **regular** on $\mathcal{R}$ in the sense that if E is in $\mathcal{R}$ and if $\epsilon > 0$ is given, then there exist a compact set K in $\mathcal{R}$ and an open set U in $\mathcal{R}$ such that $K \subseteq E \subseteq U$, $m(K) \geq m(E) - \epsilon$, and $m(U) \leq m(E) + \epsilon$: In the special case that E is a single bounded interval with endpoints a and b, we can prove regularity by taking $U = (a - \epsilon/2, b + \epsilon/2)$ and by letting $K = \varnothing$ if $b - a \leq \epsilon$ or $K = [a + \epsilon/2, b - \epsilon/2]$ if $b - a > \epsilon$. In the general case that E is the union of n bounded intervals I_j, choose K_j and U_j for I_j and for the number ϵ/n, and put $K = \bigcup_{j=1}^n K_j$ and $U = \bigcup_{j=1}^n U_j$. Then $m(K) = \sum_{j=1}^n m(K_j) \geq \sum_{j=1}^n \big(m(I_j) - \epsilon/n\big) = m(E) - \epsilon$, and Proposition 5.1f gives $m(U) \leq \sum_{j=1}^n m(U_j) \leq \sum_{j=1}^n \big(m(I_j) + \epsilon/n\big) = m(E) + \epsilon$.

Proposition 5.4. Lebesgue measure m is completely additive on the ring $\mathcal{R}$ of elementary sets in $\mathbb{R}^1$.

PROOF. Let $\{E_n\}$ be a disjoint sequence in $\mathcal{R}$ with union E in $\mathcal{R}$. Since m is nonnegative and additive, Proposition 5.1 gives $m(E) \geq m\big(\bigcup_{k=1}^n E_k\big) = \sum_{k=1}^n m(E_k)$ for every n. Passing to the limit, we obtain $m(E) \geq \sum_{k=1}^\infty m(E_k)$. For the reverse inequality, let $\epsilon > 0$ be given. Choose by regularity a compact member K of $\mathcal{R}$ and open members U_n of $\mathcal{R}$ such that $K \subseteq E$, $U_n \supseteq E_n$ for all n, $m(K) \geq m(E) - \epsilon$, and $m(U_n) \leq m(E_n) + \epsilon/2^n$. Then $K \subseteq \bigcup_{n=1}^\infty U_n$, and the compactness implies that $K \subseteq \bigcup_{n=1}^N U_n$ for some N. Hence $m(E) - \epsilon \leq m(K) \leq \sum_{n=1}^N m(U_n) \leq \sum_{n=1}^N (m(E_n) + \epsilon/2^n) \leq \sum_{n=1}^\infty m(E_n) + \epsilon$. Since ϵ is arbitrary, $m(E) \leq \sum_{n=1}^\infty m(E_n)$, and the proposition follows.

The smallest σ-ring containing the ring $\mathcal{R}$ of elementary sets in $\mathbb{R}^1$ is in fact a σ-algebra, since $\mathbb{R}^1$ is the countable union of bounded intervals. For Lebesgue measure to be truly useful, it must be extended from $\mathbb{R}$ to this σ-algebra, whose members are called the **Borel sets** of $\mathbb{R}^1$. Borel sets of $\mathbb{R}^1$ can be fairly complicated. Each open set is a Borel set because it is the countable union of bounded open intervals. Each closed set is a Borel set, being the complement of an open set, and each compact set is a Borel set because compact subsets of $\mathbb{R}^1$ are closed. In addition, any countable set, such as the set $\mathbb{Q}$ of rationals, is a Borel set as the countable union of one-point sets.

The extension is carried out by the general Extension Theorem that will be stated now and will be proved in Section 5. The theorem gives both existence and uniqueness for an extension, but not without an additional hypothesis. The need for an additional hypothesis to ensure uniqueness is closely related to the need to assume some finiteness condition on ρ in Corollary 5.3: even though each member of a decreasing sequence of sets has infinite measure, the intersection of the sets need not have infinite measure. To see what can go wrong for the Extension Theorem, consider the ring $\mathcal{R}'$ of subsets of $\mathbb{R}^1$ consisting of all finite unions of bounded intervals with rational endpoints; the individual intervals may or may not contain their endpoints. If a set function μ is defined on this ring by assigning to each set the number of elements in the set, then μ is completely additive. Each interval in $\mathbb{R}^1$ can be obtained as the union of two sets—a countable union of intervals with rational endpoints and a countable intersection of intervals with rational endpoints. It follows that the smallest σ-ring containing $\mathcal{R}'$ is the σ-algebra of all Borel sets. The set function μ can be extended to the Borel sets in more than way. In fact, each one-point set consisting of a rational must get measure 1, but a one-point set consisting of an irrational can be assigned any measure.

The additional hypothesis for the Extension Theorem is that the given nonnegative completely additive set function ν on a ring of sets $\mathcal{R}$ be σ**-finite**, i.e., that any member of $\mathcal{R}$ be contained in the countable union of members of $\mathcal{R}$ on which ν is finite. An obvious sufficient condition for σ-finiteness is that $\nu(E)$ be finite for every set in $\mathcal{R}$. This sufficient condition is satisfied by Lebesgue measure on the elementary sets, and thus the theorem proves that Lebesgue measure extends in a unique fashion to be a measure on the Borel sets.

The condition of σ-finiteness is less restrictive than a requirement that X be the countable union of sets in $\mathcal{R}$ of finite measure, another condition that is satisfied in the case of Lebesgue measure. The condition of σ-finiteness on a ring allows for some very large measures when all the sets are in a sense generated by the sets of finite measure. For example, if $\mathcal{R}$ is the ring of finite subsets of an uncountable set and ν is the counting measure, the σ-finiteness condition is satisfied. In most areas of mathematics, these very large measures rarely arise.

Theorem 5.5 (Extension Theorem). Let $\mathcal{R}$ be a ring of subsets of a nonempty set X, and let ν be a nonnegative completely additive set function on $\mathcal{R}$ that is σ-finite on $\mathcal{R}$. Then ν extends uniquely to a measure μ on the smallest σ-ring containing $\mathcal{R}$.

A **measure space** is defined to be a triple $(X, \mathcal{A}, \mu)$, where X is a nonempty set, $\mathcal{A}$ is a σ-algebra of subsets of X, and μ is a measure on X. The measure space is **finite** if $\mu(X) < +\infty$; it is σ**-finite** if X is the countable union of sets on which μ is finite. The real line, together with the σ-algebra of Borel sets and Lebesgue measure, is a σ-finite measure space.

2. Measurable Functions

In this section, X denotes a nonempty set, and $\mathcal{A}$ is a σ-algebra of subsets of X. The **measurable sets** are the members of $\mathcal{A}$.

We say that a function $f : X \to \mathbb{R}^*$ is **measurable** if

(i) $f^{-1}([-\infty, c))$ is a measurable set for every real number c.

Equivalently the measurability of f may be defined by any of the following conditions:

(ii) $f^{-1}([-\infty, c])$ is a measurable set for every real number c,
(iii) $f^{-1}((c, +\infty])$ is a measurable set for every real number c,
(iv) $f^{-1}([c, +\infty])$ is a measurable set for every real number c.

In fact, the implications (i) implies (ii), (ii) implies (iii), (iii) implies (iv), and (iv) implies (i) follow from the identities[4]

$$
\begin{aligned}
f^{-1}([-\infty, c]) &= \bigcap_{n=1}^{\infty} f^{-1}([-\infty, c + \tfrac{1}{n})),\\
f^{-1}((c, +\infty]) &= (f^{-1}([-\infty, -c]))^c,\\
f^{-1}([c, +\infty]) &= \bigcap_{n=1}^{\infty} f^{-1}((c - \tfrac{1}{n}, +\infty]),\\
f^{-1}([-\infty, c)) &= (f^{-1}([-c, +\infty]))^c.
\end{aligned}
$$

EXAMPLES.

(1) If $\mathcal{A} = \{\varnothing, X\}$, then only the constant functions are measurable.

(2) If $\mathcal{A}$ consists of all subsets of X, then every function from X to $\mathbb{R}^*$ is measurable.

[4] Manipulations with inverse images of sets are discussed in Section A1 of the appendix.

(3) If $X = \mathbb{R}^1$ and $\mathcal{A}$ consists of the Borel sets of $\mathbb{R}^1$, the measurable functions are often called **Borel measurable**. Every continuous function is Borel measurable by (i) because the inverse image of every open set is open. Any function that is 1 on an open or compact set and is 0 off that set is Borel measurable. It is shown in Problem 11 at the end of the chapter that not every Riemann integrable function (when set equal to 0 off some bounded interval) is Borel measurable. However, let us verify that every function that is continuous except at countably many points is Borel measurable. In fact, let C be the exceptional countable set. The restriction of f to the metric space $\mathbb{R}-C$ is continuous, and hence the inverse image in $\mathbb{R} - C$ of any open set $[-\infty, c)$ is open in $\mathbb{R} - C$. Hence the inverse image is the countable union of sets $(a, b) - C$, and these are Borel sets. The full inverse image in $\mathbb{R}$ of $[-\infty, c)$ under f is the union of a countable set and this subset of $\mathbb{R} - C$ and hence is a Borel set.

(4) If $X = \mathbb{R}^1$ and if $\mathcal{A}$ consists of the "Lebesgue measurable sets" in a sense to be defined in Section 5, the measurable functions are often called **Lebesgue measurable**. Every Borel measurable function is Lebesgue measurable, and so is every Riemann integrable function (when set equal to 0 off some bounded interval).

The next proposition discusses, among other things, functions f^+, f^-, and $|f|$ defined by $f^+(x) = \max\{f(x), 0\}$, $f^-(x) = -\min\{f(x), 0\}$, and $|f|(x) = |f(x)|$. Then $f = f^+ - f^-$ and $|f| = f^+ + f^-$.

Proposition 5.6.

(a) Constant functions are always measurable.

(b) If f is measurable, then the inverse image of any interval is measurable.

(c) If f is measurable, then the inverse image of any open set in $\mathbb{R}^*$ is measurable.

(d) If f is measurable, then the functions f^+, f^-, and $|f|$ are measurable.

PROOF. In (a), the inverse image of a set under a constant function is either $\varnothing$ or X and in either case is measurable. In (b), the inverse image of an interval is the intersection of two sets of the kind described in (i) through (iv) above and hence is measurable. In (c), any open set in $\mathbb{R}^*$ is the countable union of open intervals, and the measurability of the inverse image follows from (b) and the closure of $\mathcal{A}$ under countable unions. In (d), $(f^+)^{-1}((c, +\infty))$ equals $f^{-1}((c, +\infty))$ if $c \geq 0$ and equals X if $c < 0$. The measurability of f^- and $|f|$ are handled similarly.

Next we deal with measurability of sums and products, allowing for values $+\infty$ and $-\infty$. Recall from Section I.1 that multiplication is everywhere defined in $\mathbb{R}^*$ and that the product in $\mathbb{R}^*$ of 0 with anything is 0.

Proposition 5.7. Let f and g be measurable functions, and let a be in $\mathbb{R}$. Then af and fg are measurable, and $f+g$ is measurable provided the sum $f(x)+g(x)$ is everywhere defined.

PROOF. For $f+g$, with $\mathbb{Q}$ denoting the rationals,

$$(f+g)^{-1}(c,+\infty] = \bigcup_{r\in\mathbb{Q}} f^{-1}(c+r,+\infty] \cap g^{-1}(-r,+\infty].$$

If $a=0$, then $af=0$, and 0 is measurable. If $a\neq 0$, then

$$(af)^{-1}(c,+\infty] = \begin{cases} f^{-1}\big(\frac{c}{a},+\infty\big] & \text{if } a>0, \\ f^{-1}\big[-\infty,\frac{c}{a}\big) & \text{if } a<0. \end{cases}$$

If f and g are measurable and are ≥ 0, then

$$(fg)^{-1}(c,+\infty] = \begin{cases} \bigcup_{r\in\mathbb{Q},\, r>0} f^{-1}\big(\frac{c}{r},+\infty\big] \cap g^{-1}(r,+\infty] & \text{if } c\geq 0, \\ X & \text{if } c<0. \end{cases}$$

Hence fg is measurable in this special case. In the general case the formula $fg = f^+g^+ + f^-g^- - f^+g^- - f^-g^+$ exhibits fg as the everywhere-defined sum of measurable functions.

Proposition 5.8. If $\{f_n\}$ is a sequence of measurable functions, then the functions

(a) $\sup_n f_n$,
(b) $\inf_n f_n$,
(c) $\limsup_n f_n$,
(d) $\liminf f_n$,

are all measurable.

PROOF. For (a) and (b), we have $(\sup f_n)^{-1}(c,+\infty] = \bigcup_{n=1}^{\infty} f_n^{-1}(c,+\infty]$ and $(\inf f_n)^{-1}([-\infty,c) = \bigcup_{n=1}^{\infty} f_n^{-1}[-\infty,c)$. For (c) and (d), we have $\limsup_n f_n = \inf_n \sup_{k\geq n} f_k$ and $\liminf_n f_n = \sup_n \inf_{k\geq n} f_k$.

Corollary 5.9. The pointwise maximum and the pointwise minimum of a finite set of measurable functions are both measurable.

PROOF. These are special cases of (a) and (b) in the proposition.

Corollary 5.10. If $\{f_n\}$ is a sequence of measurable functions and if $f(x) = \lim f_n(x)$ exists in $\mathbb{R}^*$ at every x, then f is measurable.

PROOF. This is the special case of (c) and (d) in the proposition in which $\limsup_n f_n = \liminf_n f_n$.

The above results show that the set of measurable functions is closed under pointwise limits, as well as the arithmetic operations and max and min. Since the measurable functions will be the ones we attempt to integrate, we can hope for good limit theorems from Lebesgue integration, as well as the familiar results about arithmetic operations and ordering properties.

If E is a subset of X, the **indicator function**[5] I_E of E is the function that is 1 on E and is 0 elsewhere. The set $(I_E)^{-1}(c, +\infty]$ is $\varnothing$ or E or X, depending on the value of c. Therefore I_E is a measurable function if and only if E is a measurable set.

A **simple function** $s : X \to \mathbb{R}^*$ is a function s with finite image contained in $\mathbb{R}$. Every simple function s has a unique representation as $s = \sum_{n=1}^N c_n I_{E_n}$, where the c_n are distinct real numbers and the E_n are disjoint nonempty sets with union X. In fact, the set of numbers c_n equals the image of s, and E_n is the set where s takes the value c_n. This expansion of s will be called the **canonical expansion** of s. The set $s^{-1}(c, +\infty]$ is the union of the sets E_n such that $c < c_n$, and it follows that s is a measurable function if and only if all of the sets E_n in the canonical expansion are measurable sets.

Proposition 5.11. For any function $f : X \to [0, +\infty]$, there exists a sequence of simple functions $s_n \geq 0$ with the property that for each x in X, $\{s_n(x)\}$ is a monotone increasing sequence in $\mathbb{R}$ with limit $f(x)$ in $\mathbb{R}^*$. If f is measurable, then the simple functions s may be taken to be measurable.

PROOF. For $1 \leq n < \infty$ and $1 \leq j \leq n2^n$, let

$$E_{nj} = f^{-1}\Big[\frac{j-1}{2^n}, \frac{j}{2^n}\Big), \quad F_n = f^{-1}[n, +\infty), \quad s_n = \sum_{j=1}^{n2^n} \frac{j-1}{2^n} I_{E_{nj}} + n I_{F_n}.$$

Then $\{s_n\}$ has the required properties.

By convention from now on, *simple functions will always be understood to be measurable*.

3. Lebesgue Integral

Throughout this section, $(X, \mathcal{A}, \mu)$ denotes a measure space. The measurable sets continue to be those in $\mathcal{A}$. Our objective in this section is to define the Lebesgue

[5] As noted in Chapter III, indicator functions are called "characteristic functions" by many authors, but the term "characteristic function" has another meaning in probability theory and is best avoided as a substitute for "indicator function" in any context where probability might play a role.

integral. We defer any systematic discussion of properties of the integral to Section 4.

Just as with the Riemann integral, the Lebesgue integral is defined by means of an approximation process. In the case of the Riemann integral, the process is to use upper sums and lower sums, which capture an approximate value of an integral by adding contributions influenced by proximity in the *domain* of the integrand. The process is qualitatively different for the Lebesgue integral, which captures an approximate value of an integral by adding contributions based on what happens in the *image* of the integrand.

Let s be a simple function ≥ 0. By our convention at the end of the previous section, we have incorporated measurability into the definition of simple function. Let E be a measurable set, and let $s = \sum_{n=1}^{N} c_n I_{A_n}$ be the canonical expansion of s. We define $\mathcal{I}_E(s) = \sum_{n=1}^{N} c_n \mu(A_n \cap E)$. This kind of object will be what we use as an approximation in the definition of the Lebesgue integral; the formula shows the sense in which $\mathcal{I}_E(s)$ is built from the *image* of the integrand.

If $f \geq 0$ is a measurable function and E is a measurable set, we define the **Lebesgue integral** of f on the set E with respect to the measure μ to be

$$\int_E f\, d\mu = \int_E f(x)\, d\mu(x) = \sup_{\substack{0 \leq s \leq f, \\ s \text{ simple}}} \mathcal{I}_E(s).$$

This is well-defined as a member of $\mathbb{R}^*$ without restriction as long as E is a measurable set and the measurable function f is ≥ 0 everywhere on X. It is evident in this case that $\int_E f\, d\mu \geq 0$ and that $\int_E 0\, d\mu = 0$.

For a general measurable function f, not necessarily ≥ 0, the integral may or may not be defined. We write $f = f^+ - f^-$. The functions f^+ and f^- are ≥ 0 and are measurable by Proposition 5.6d, and consequently $\int_E f^+\, d\mu$ and $\int_E f^-\, d\mu$ are well-defined members of $\mathbb{R}^*$. If $\int_E f^+\, d\mu$ and $\int_E f^-\, d\mu$ are not both infinite, then we define

$$\int_E f\, d\mu = \int_E f(x)\, d\mu(x) = \int_E f^+\, d\mu - \int_E f^-\, d\mu.$$

This definition is consistent with the definition in the special case $f \geq 0$, since such an f has $f^- = 0$ and therefore $\int_E f^-\, d\mu = 0$. We say that f is **integrable** if $\int_E f^+\, d\mu$ and $\int_E f^-\, d\mu$ are *both* finite. In this case the subsets of E where f is $+\infty$ and where f is $-\infty$ have measure 0. In fact, if S is the subset of E where f^+ is $+\infty$, then the inequality $\int_E f^+\, d\mu \geq \mathcal{I}_E(C I_S) = C\mu(S)$ for every $C > 0$ shows that $\mu(S) \leq C^{-1} \int_E f^+\, d\mu$ for every C; hence $\mu(S) = 0$. A similar argument applies to the set where f^- is $+\infty$.

We shall give some examples of integration after showing that the definition of $\int_E f\, d\mu$ reduces to $\mathcal{I}_E(f)$ if f is nonnegative and simple. The first lemma below will make use of the additivity of μ, and the second lemma will make use of the fact that μ is nonnegative.

Lemma 5.12. Let $s = \sum_{n=1}^{N} c_n I_{A_n}$ be the canonical expansion of a simple function ≥ 0, and let $s = \sum_{m=1}^{M} d_n I_{B_m}$ be another expansion in which the d_m are ≥ 0 and the B_m are disjoint and measurable. Then $\mathcal{I}_E(s) = \sum_{m=1}^{M} d_m \mu(B_m \cap E)$.

PROOF. Adjoin the term $0 \cdot I_{(\bigcup_m B_m)^c}$ to the second expansion, if necessary, to make $\bigcup_{m=1}^{M} B_m = X$. Without loss of generality, we may assume that no B_m is empty. Then the fact that the sets B_m are disjoint and nonempty with union X implies that the image of s is $\{d_1, \dots, d_M\}$. Thus we can write $d_m = c_{n(m)}$ for each m. Since $A_n = s^{-1}(\{c_n\})$, we see that $B_m \subseteq A_{n(m)}$. Since the B_m are disjoint with union X, we obtain

$$A_k = \bigcup_{\{m \mid n(m)=k\}} B_m$$

disjointly. The additivity of μ gives $\mu(A_k \cap E) = \sum_{\{m \mid n(m)=k\}} \mu(B_m \cap E)$, and thus $c_k \mu(A_k \cap E) = \sum_{\{m \mid n(m)=k\}} d_m \mu(B_m \cap E)$. Summing on k, we obtain the conclusion of the lemma.

Lemma 5.13. If s and t are nonnegative simple functions and if $t \leq s$ on E, then $\mathcal{I}_E(t) \leq \mathcal{I}_E(s)$.

PROOF. If $s = \sum_{j=1}^{J} c_j I_{A_j}$ and $t = \sum_{k=1}^{K} d_k I_{B_k}$ are the canonical expansions of s and t, then $\bigcup_{j,k} (A_j \cap B_k) = X$ disjointly. Hence we can write

$$s = \sum_{j,k} c_j I_{A_j \cap B_k} \qquad \text{and} \qquad t = \sum_{j,k} d_k I_{A_j \cap B_k}.$$

Lemma 5.12 shows that

$$\mathcal{I}_E(s) = \sum_{j,k} c_j \mu(A_j \cap B_k \cap E) \qquad \text{and} \qquad \mathcal{I}_E(t) = \sum_{j,k} d_k \mu(A_j \cap B_k \cap E).$$

We now have term-by-term inequality: either $\mu(A_j \cap B_k \cap E) = 0$ for a term, or $A_j \cap B_k \cap E \neq \varnothing$ and any x in $A_j \cap B_k \cap E$ has $t(x) \leq s(x)$ and exhibits $d_k \leq c_j$.

Proposition 5.14. If $s \geq 0$ is a simple function, then $\int_E s\,d\mu = \mathcal{I}_E(s)$ for every measurable set E.

PROOF. If t is a simple function with $0 \leq t \leq s$ everywhere, then Lemma 5.13 gives $\mathcal{I}_E(t) \leq \mathcal{I}_E(s)$. Hence $\int_E s\,d\mu = \sup_{0 \leq t \leq s} \mathcal{I}_E(t) \leq \mathcal{I}_E(s)$. On the other hand, we certainly have $\mathcal{I}_E(s) \leq \sup_{0 \leq t \leq s} \mathcal{I}_E(t) = \int_E s\,d\mu$, and thus $\int_E s\,d\mu = \mathcal{I}_E(s)$.

EXAMPLES.

(1) Let $\mathcal{A} = \{\varnothing, X\}$ and $\mu(X) = 1$. Only the constant functions are measurable, and $\int_\varnothing c\,d\mu = 0$ and $\int_X c\,d\mu = c$.

(2) Let X be a nonempty countable set, let $\mathcal{A}$ consist of all subsets of X, and let μ be defined by nonnegative finite weights w_i attached to each point i in X. If $f = \{f_i\}$ is a real-valued function, then the integral of f over X is $\sum f_i w_i$ provided the integrals of f^+ and f^- are not both infinite, i.e., provided every rearrangement of the series $\sum f_i w_i$ converges in $\mathbb{R}^*$ to the same sum. By contrast, f is integrable if and only if the series $\sum f_i w_i$ is absolutely convergent. In the special case that all the weights w_i are 1, the theory of the Lebesgue integral over X reduces to the theory of infinite series for which every rearrangement of the series converges in $\mathbb{R}^*$ to the same sum. This is a very important special case for testing the validity of general assertions about Lebesgue integration.

(3) Let $(X, \mathcal{A}, \mu)$ be the real line $\mathbb{R}^1$ with $\mathcal{A}$ consisting of the Borel sets and with μ equal to Lebesgue measure m. Recall that real-valued continuous functions on $\mathbb{R}^1$ are measurable. For such a function f, the assertion is that

$$\int_{[a,x)} f\,dm = \int_a^x f(t)\,dt,$$

the left side being a Lebesgue integral and the right side being a Riemann integral. Proving this assertion involves using some properties of the Lebesgue integral that will be proved in the next section. We give the argument now before these properties have been established, in order to emphasize the importance of each of these properties: If $h > 0$, then

$$\begin{aligned}\frac{1}{h}\Big[\int_{[a,x+h)} f\,dm - \int_{[a,x)} f\,dm\Big] - f(x) &= \frac{1}{h}\int_{[x,x+h)} f\,dm - f(x)\\ &= \frac{1}{h}\int_{[x,x+h)} [f - f(x)]\,dm.\end{aligned}$$

The absolute value of the left side is then

$$\begin{aligned}\le \frac{1}{h}\int_{[x,x+h)} |f - f(x)|\,dm &\le \frac{1}{h}\sup_{t\in[x,x+h)} |f(t) - f(x)|\,m([x, x+h))\\ &= \sup_{t\in[x,x+h)} |f(t) - f(x)|,\end{aligned}$$

and the right side tends to 0 as h decreases to 0, by continuity of f at x. If $h < 0$, then the argument corresponding to the first display is

$$\begin{aligned}\frac{1}{h}\Big[\int_{[a,x+h)} f\,dm - \int_{[a,x)} f\,dm\Big] - f(x) &= -\frac{1}{h}\int_{[x-|h|,x)} f\,dm - f(x)\\ &= \frac{1}{|h|}\int_{[x-|h|,x)} [f - f(x)]\,dm.\end{aligned}$$

The absolute value of the left side is then $\le \sup_{t\in[x-|h|,x)} |f(t) - f(x)|$, and this tends to 0 as h increases to 0, by continuity of f at x. We conclude that $\int_{[a,\cdot)} f\,dm$ is differentiable with derivative f. By the Fundamental Theorem of Calculus for the Riemann integral, together with a corollary of the Mean Value Theorem, $\int_{[a,x)} f\,dm = \int_a^x f(t)\,dt + c$ for all x and some constant c. Putting $x = a$, we see that $c = 0$. Therefore the Riemann and Lebesgue integrals coincide for continuous functions on bounded intervals $[a, b)$.

4. Properties of the Integral

In this section, $(X, \mathcal{A}, \mu)$ continues to denote a measure space. Our objective is to establish basic properties of the Lebesgue integral, including properties that indicate how Lebesgue integration interacts with passages to the limit. The properties that we establish will include all remaining properties needed to justify the argument in Example 3 at the end of the previous section.

Proposition 5.15. The Lebesgue integral has these four properties:

(a) If f is a measurable function and $\mu(E) = 0$, then $\int_E f\,d\mu = 0$.

(b) If E and F are measurable sets with $F \subseteq E$ and if f is a measurable function, then $\int_F f^+\,d\mu \le \int_E f^+\,d\mu$ and $\int_F f^-\,d\mu \le \int_E f^-\,d\mu$. Consequently, if $\int_E f\,d\mu$ is defined, then so is $\int_F\,d\mu$.

(c) If c is a constant function with its value in $\mathbb{R}^*$, then $\int_E c\,d\mu = c\mu(E)$.

(d) If $\int_E f\,d\mu$ is defined and if c is in $\mathbb{R}$, then $\int_E cf\,d\mu$ is defined and $\int_E cf\,d\mu = c\int_E f\,d\mu$. If f is integrable on E, then so is cf.

PROOF. In (a), it is enough to deal with f^+ and f^- separately, and then it is enough to handle $s \ge 0$ simple. For such an s, Proposition 5.14 says that the integral equals $\mathcal{I}_E(s)$, and the definition shows that this is 0. In (b), Proposition 5.14 makes it clear that the inequalities are valid for any simple function ≥ 0, and then the general case follows by taking the supremum first for $0 \le s \le f^+$ and then for $0 \le s \le f^-$. In (c), if $0 \le c < +\infty$, then c is simple, and the integral equals $\mathcal{I}_E(c) = c\mu(E)$ by Proposition 5.14. If $c = +\infty$, then the case $\mu(E) = 0$ follows from (a) and the case $\mu(E) > 0$ is handled by the observations that $\int_E c\,d\mu \ge \mathcal{I}_E(n) = n\mu(E)$ and that the right side tends to $+\infty$ as n tends to $+\infty$. For $c \le 0$, we have $\int_E c\,d\mu = -\int_E (-c)\,d\mu$ by definition, and then the result follows from the previous cases. In (d), we may assume, without loss of generality, that $f \ge 0$ and $c \ge 0$. Then $\int_E cf\,d\mu = \sup_{0\le s\le cf} \mathcal{I}_E(s) = \sup_{0\le ct\le cf} \mathcal{I}_E(ct) = c \sup_{0\le t\le f} \mathcal{I}_E(t) = c\int_E f\,d\mu$, and (d) is proved.

Proposition 5.16. If f and g are measurable functions, if their integrals over E are defined, and if $f(x) \le g(x)$ on E, then $\int_E f\,d\mu \le \int_E g\,d\mu$.

REMARK. Observe that the inequality $f(x) \le g(x)$ is assumed only on E, despite the definitions that take into account values of a function everywhere on X. This "localization" property of the integral is as one wants it to be.

PROOF. First suppose that $f \ge 0$ and $g \ge 0$. If s is any simple function with $0 \le s \le f$, define t to equal s on E and to equal 0 off E. Then $0 \le t \le g$, and Lemma 5.13 gives $\mathcal{I}_E(s) = \mathcal{I}_E(t) \le \int_E g\,d\mu$. Hence $\int_E f\,d\mu \le \int_E g\,d\mu$ when $f \ge 0$ and $g \ge 0$.

In the general case the inequality $f(x) \le g(x)$ on E implies that $f^+(x) \le g^+(x)$ on E and $f^-(x) \ge g^-(E)$ on E. The special case gives $\int_E f^+\,d\mu \le \int_E g^+\,d\mu$ and $\int_E f^-\,d\mu \ge \int_E g^-\,d\mu$. Subtracting these inequalities, we obtain the desired result.

Corollary 5.17. If f and g are measurable functions that are equal on E and if $\int_E f\,d\mu$ is defined, then $\int_E g\,d\mu$ is defined and $\int_E f\,d\mu = \int_E g\,d\mu$.

PROOF. Apply Proposition 5.16 to the following inequalities on E, and then sort out the results: $f^+ \le g^+$, $f^+ \ge g^+$, $f^- \le g^-$, and $f^- \ge g^-$.

Corollary 5.18. If f is a measurable function, then f is integrable on E if either

(a) there is a function g integrable on E such that $|f(x)| \le g(x)$ on E, or
(b) $\mu(E)$ is finite and there is a real number c such that $|f(x)| \le c$ on E.

PROOF. For (a), apply Proposition 5.16 to the inequalities $f^+ \le g$ and $f^- \le g$ valid on E. For (b), use the formula for $\int_E c\,d\mu$ in Proposition 5.15c and apply (a).

We turn our attention now to properties that indicate how Lebesgue integration interacts with passages to the limit. These make essential use of the complete additivity of the measure μ. We shall bring this hypothesis to bear initially through the following theorem.

Theorem 5.19. Let f be a fixed measurable function, and suppose that $\int_X f\,d\mu$ is defined. Then the set function $\rho(E) = \int_E f\,d\mu$ is completely additive.

PROOF. We have $\rho(\varnothing) = 0$ by Proposition 5.15a, since $\mu(\varnothing) = 0$. We shall prove that if $f \ge 0$, then ρ is completely additive. The general case follows from this by applying the result to f^+ and f^- separately and by using the fact that $\int_X f^+\,d\mu$ and $\int_X f^-\,d\mu$ are not both infinite. Thus we are to show that if $E = \bigcup_{n=1}^\infty E_n$ disjointly and if $f \ge 0$, then $\rho(E) = \sum_{n=1}^\infty \rho(E_n)$.

For simple $s \ge 0$ with canonical expansion $s = \sum_{n=1}^N c_n I_{A_n}$, the identity $\mathcal{I}_F(s) = \sum_{n=1}^N c_n \mu(A_n \cap F)$ and the complete additivity of μ show that $\mathcal{I}_F(s)$ is a completely additive function of the set F. Thus for s simple with $0 \le s \le f$, we have

$$\mathcal{I}_E(s) = \sum_{n=1}^{\infty} \mathcal{I}_{E_n}(s) \le \sum_{n=1}^{\infty} \rho(E_n).$$

Hence
$$\rho(E) = \sup_{0 \le s \le f} \mathcal{I}_E(s) \le \sum_{n=1}^{\infty} \rho(E_n).$$

We now prove the reverse inequality. By Proposition 5.15b, $\rho(E) \ge \rho(E_n)$ for every n, since $f = f^+$. Hence if $\rho(E_n) = +\infty$ for any n, the desired result is proved. Thus assume that $\rho(E_n) < +\infty$ for all n. Let $\epsilon > 0$ be given, and choose simple functions t and u that are ≥ 0 and are $\le f$ and have

$$\mathcal{I}_{E_1}(t) \ge \int_{E_1} f\,d\mu - \epsilon \qquad \text{and} \qquad \mathcal{I}_{E_2}(u) \ge \int_{E_2} f\,d\mu - \epsilon.$$

Let s be the pointwise maximum $s = \max\{t, u\}$. Then s is simple, and Lemma 5.13 gives $\mathcal{I}_{E_1}(s) \ge \mathcal{I}_{E_1}(t)$ and $\mathcal{I}_{E_2}(s) \ge \mathcal{I}_{E_2}(u)$. Consequently

$$\begin{aligned}
\rho(E_1 \cup E_2) &= \int_{E_1 \cup E_2} f\,d\mu \ge \mathcal{I}_{E_1 \cup E_2}(s) = \mathcal{I}_{E_1}(s) + \mathcal{I}_{E_2}(s) \\
&\ge \mathcal{I}_{E_1}(t) + \mathcal{I}_{E_2}(u) \ge \int_{E_1} f\,d\mu + \int_{E_2} f\,d\mu - 2\epsilon \\
&= \rho(E_1) + \rho(E_2) - 2\epsilon.
\end{aligned}$$

Since ϵ is arbitrary, $\rho(E_1 \cup E_2) \ge \rho(E_1) + \rho(E_2)$. By induction, we obtain $\rho(E_1 \cup \cdots \cup E_n) \ge \rho(E_1) + \cdots + \rho(E_n)$ for every n, and thus $\rho(E) \ge \rho(E_1) + \cdots + \rho(E_n)$ by another application of Proposition 5.15b. Therefore $\rho(E) \ge \sum_{n=1}^{\infty} \rho(E_n)$, and the reverse inequality has been proved.

We give five corollaries that are consequences of Corollary 5.17 and Theorem 5.19. The first three make use only of additivity, not of complete additivity.

Corollary 5.20. If $\int_E f\,d\mu$ is defined, then $\int_X I_E f\,d\mu$ is defined and equals $\int_E f\,d\mu$.

PROOF. It is sufficient to handle f^+ and f^- separately. Then both integrals are defined, and $\int_E f\,d\mu = \int_E I_E f\,d\mu + \int_{E^c} 0\,d\mu = \int_E I_E f\,d\mu + \int_{E^c} I_E f\,d\mu = \int_X I_E f\,d\mu$.

Corollary 5.21. If $\int_E f\,d\mu$ is defined, then $\big|\int_E f\,d\mu\big| \le \int_E |f|\,d\mu$. If f is integrable on E, so is $|f|$.

PROOF. Let $E_1 = E \cap f^{-1}([0,+\infty])$ and $E_2 = E \cap f^{-1}([-\infty,0))$. Then use of the triangle inequality gives

$$\begin{aligned}\big|\textstyle\int_E f\,d\mu\big| &= \big|\textstyle\int_{E_1} f^+\,d\mu - \int_{E_2} f^-\,d\mu\big| \le \textstyle\int_{E_1} f^+\,d\mu + \int_{E_2} f^-\,d\mu \\ &= \textstyle\int_{E_1} |f|\,d\mu + \int_{E_2} |f|\,d\mu = \int_E |f|\,d\mu.\end{aligned}$$

If f is integrable on E, both $\int_{E_1} f^+\,d\mu$ and $\int_{E_2} f^-\,d\mu$ are finite. Their sum is $\int_E |f|\,d\mu$.

Corollary 5.22. If f is a measurable function and $\mu(E \,\Delta\, F) = 0$, then $\int_E f\,d\mu = \int_F f\,d\mu$, provided one of the integrals exists.

PROOF. Without loss of generality, we may assume that $f \ge 0$. Then both integrals are defined. Since $E \,\Delta\, F = (E-F)\cup(F-E)$, we have $\mu(E-F) = \mu(F-E) = 0$. Then Theorem 5.19 and Proposition 5.15a give $\int_E f\,d\mu = \int_{E-F} f\,d\mu + \int_{E\cap F} f\,d\mu = 0 + \int_{E\cap F} f\,d\mu = \int_{F-E} f\,d\mu + \int_{E\cap F} f\,d\mu = \int_F f\,d\mu$.

Corollary 5.23. If f is a measurable function and if the set $A = \{x \mid f(x) \ne 0\}$ has $\mu(A) = 0$, then $\int_X f\,d\mu = 0$. Conversely if f is measurable, is ≥ 0, and has $\int_X f\,d\mu = 0$, then $A = \{x \mid f(x) \ne 0\}$ has $\mu(A) = 0$.

REMARKS. When a set where some condition fails to hold has measure 0, one sometimes says that the condition holds **almost everywhere**, or **a.e.**, or at **almost every** point. If there is any ambiguity about what measure is being referred to, one says "a.e. $[d\mu]$." Thus the conclusion in the converse half of the above proposition is that f is zero a.e. $[d\mu]$.

PROOF. For the first statement, Corollary 5.20 gives $\int_X f\,d\mu = \int_X I_A f\,d\mu = \int_A f\,d\mu = 0$. Conversely let $A_n = f^{-1}\big(\big[\tfrac{1}{n},+\infty\big]\big)$. This is a measurable set. Since f is ≥ 0, $A = \bigcup_{n=1}^\infty A_n$. Proposition 5.1g and complete additivity of μ give $\mu(A) \le \sum_{n=1}^\infty \mu(A_n)$. If $\mu(A_n) > 0$ for some n, then $\int_X f\,d\mu = \int_{A_n} f\,d\mu + \int_{A_n^c} f\,d\mu \ge \int_{A_n} \frac{1}{n}\,d\mu = \frac{1}{n}\,\mu(A_n) > 0$, and we obtain a contradiction. We conclude that $\mu(A_n) = 0$ for all n and hence that $\mu(A) = 0$.

Corollary 5.24. If $f \ge 0$ is an integrable function on X, then for any $\epsilon > 0$, there exists a $\delta > 0$ such that $\int_E f\,d\mu \le \epsilon$ for every measurable set E with $\mu(E) \le \delta$.

PROOF. Let $\epsilon > 0$ be given. If $N > 0$ is an integer, then the sets $S_N = \{x \in X \mid f(x) \geq N\}$ form a decreasing sequence whose intersection is $S = \{x \in X \mid f(x) = +\infty\}$. Since f is integrable, $\mu(S) = 0$ and therefore $\int_S f\,d\mu = 0$. The finiteness of $\int_X f\,d\mu$, together with Corollary 5.3 and the complete additivity of $E \mapsto \int_E f\,d\mu$ given in Theorem 5.19, implies that $\lim_N \int_{S_N} f\,d\mu = 0$. Choose N large enough so that $\int_{S_N} f\,d\mu \leq \epsilon/2$, and then choose $\delta = \epsilon/(2N)$. If $\mu(E) \leq \delta$, then

$$\begin{aligned}\textstyle\int_E f\,d\mu &= \textstyle\int_{S_N \cap E} f\,d\mu + \int_{S_N^c \cap E} f\,d\mu \\ &\leq \textstyle\int_{S_N} f\,d\mu + \int_{S_N^c \cap E} N\,d\mu \leq \epsilon/2 + N\mu(E) \leq \epsilon/2 + \epsilon/2 = \epsilon,\end{aligned}$$

and the proof is complete.

In a number of the remaining results in the section, a sequence $\{f_n\}$ of measurable functions converges pointwise to a function f. Corollary 5.10 assures us that f is measurable. Suppose that $\int_E f_n\,d\mu$ exists for each n. Is it true that $\int_E f\,d\mu$ exists, is it true that $\lim_n \int_E f_n\,d\mu$ exists, and if both exist, are they equal? Once again we encounter an interchange-of-limits problem, and there is no surprise from the general fact: all three answers can be "no" in particular cases. Examples of the failure of the limit of the integral to equal the integral of the limit are given below. After giving the examples, we shall discuss theorems that give "yes" answers under additional hypotheses.

EXAMPLES.

(1) Let X be the set of positive integers, let $\mathcal{A}$ consist of all subsets of X, and let μ be counting measure. A measurable function f is a sequence $\{f(k)\}$ with values in $\mathbb{R}^*$. Define a sequence $\{f_n\}$ of measurable functions for $n \geq 1$ by taking

$$f_n(k) = \begin{cases} 1/n & \text{if } k \leq n, \\ 0 & \text{if } k > n. \end{cases}$$

Then $\int_X f_n\,d\mu = 1$ for all n, $\lim f_n = 0$ pointwise, and

$$\int_X \lim f_n\,d\mu < \lim \int_X f_n\,d\mu.$$

(2) Let the measure space be $X = \mathbb{R}^1$ with the Borel sets and Lebesgue measure m. Define

$$f_n(x) = \begin{cases} n & \text{for } 0 < x < 1/n, \\ 0 & \text{otherwise.} \end{cases}$$

Then the same phenomenon results, and everything of interest is taking place within $[0, 1]$. So the difficulty in the previous example does not result from the fact that X has infinite measure.

Theorem 5.25 (Monotone Convergence Theorem). Let E be a measurable set, and suppose that $\{f_n\}$ is a sequence of measurable functions that satisfy

$$0 \le f_1(x) \le f_2(x) \le \cdots \le f_n(x) \le \cdots$$

for all x. Put $f(x) = \lim_n f_n(x)$, the limit being taken in $\mathbb{R}^*$. Then $\int_E f\,d\mu$ and $\lim_n \int_E f_n\,d\mu$ both exist, and

$$\int_E f\,d\mu = \lim_{n\to\infty} \int_E f_n\,d\mu.$$

REMARKS. This theorem generalizes Corollary 1.14, which is the special case of the Monotone Convergence Theorem in which X is the set of positive integers, every subset is measurable, and μ is counting measure. In the general setting of the Monotone Convergence Theorem, one of the by-products of the theorem is that we obtain an easier way of dealing with the definition of $\int_E f\,d\mu$ for $f \ge 0$. Instead of using the totality of simple functions between 0 and f, we may use a single increasing sequence with pointwise f, such as the one given by Proposition 5.11. The proof of Proposition 5.26 below will illustrate how we can take advantage of this fact.

PROOF. Since f is the pointwise limit of measurable functions and is ≥ 0, f is measurable and $\int_E f\,d\mu$ exists in $\mathbb{R}^*$. Since $\{f_n(x)\}$ is monotone increasing in n, the same is true of $\left\{\int_E f_n\,d\mu\right\}$. Therefore $\lim_n \int_E f_n\,d\mu$ exists in $\mathbb{R}^*$. Let us call this limit k. For each n, $\int_E f_n\,d\mu \le \int_E f\,d\mu$ because $f_n \le f$. Therefore $k \le \int_E f\,d\mu$, and the problem is to prove the reverse inequality.

Let c be any real number with $0 < c < 1$, to be regarded as close to 1, and let s be a simple function with $0 \le s \le f$. Define

$$E_n = \left\{x \in E \;\middle|\; f_n(x) \ge cs(x)\right\}.$$

These sets are measurable, and $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots \subseteq E$. Let us see that $E = \bigcup_{n=1}^\infty E_n$. If $f(x) = 0$ for a particular x in E, then $f_n(x) = 0$ for all n and also $cs(x) = 0$. Thus x is in every E_n. If $f(x) > 0$, then the inequality $f(x) \ge s(x)$ forces $f(x) > cs(x)$. Since $f_n(x)$ has increasing limit $f(x)$, $f_n(x)$ must be $> cs(x)$ eventually, and then x is in E_n. In either case x is in $\bigcup_{n=1}^\infty E_n$. Thus $E = \bigcup_{n=1}^\infty E_n$.

For every n, we have

$$k \ge \int_E f_n\,d\mu \ge \int_{E_n} f_n\,d\mu \ge \int_{E_n} cs\,d\mu = c\int_{E_n} s\,d\mu.$$

Since, by Theorem 5.19, the integral is a completely additive set function, Proposition 5.2 shows that $\lim \int_{E_n} s\,d\mu = \int_E s\,d\mu$. Therefore $k \ge c\int_E s\,d\mu$. Since c is arbitrary with $0 < c < 1$, $k \ge \int_E s\,d\mu$. Taking the supremum over s with $0 \le s \le f$, we conclude that $k \ge \int_E f\,d\mu$.

Proposition 5.26. If f and g are measurable functions, if their sum $h = f + g$ is everywhere defined, and if $\int_E f\,d\mu + \int_E g\,d\mu$ is defined, then $\int_E h\,d\mu$ is defined and

$$\int_E h\,d\mu = \int_E f\,d\mu + \int_E g\,d\mu.$$

REMARK. It may seem surprising that complete additivity plays a role in the proof of this proposition, since it apparently played no role in the linearity of the Riemann integral. In fact, although complete additivity is used when f and g are unbounded, it can be avoided when f and g are bounded, as will be observed in Problems 42–43 at the end of the chapter. The distinction between the two cases is that the pointwise convergence in Proposition 5.11 is actually uniform if the given function is bounded, whereas it cannot be uniform for an unbounded function because the uniform limit of bounded functions is bounded.

PROOF. The sum h is measurable by Proposition 5.7. For the conclusions about integration, first assume that $f \geq 0$ and $g \geq 0$. In the case of simple functions $s = t + u$ with $t \geq 0$ and $u \geq 0$, we use Proposition 5.14 and Lemma 5.12. The proposition shows that we are to prove that $\mathcal{I}_E(s) = \mathcal{I}_E(t) + \mathcal{I}_E(u)$, and the lemma shows that we can use expansions of t and u into sets on which t and u are both constant and the conclusion about $\mathcal{I}_E(s)$ is evident. If f and g are ≥ 0 and are not necessarily simple, then we can use Proposition 5.11 to find increasing sequences $\{t_n\}$ and $\{u_n\}$ of simple functions ≥ 0 with limits f and g. If $s_n = t_n + u_n$, then s_n is nonnegative simple, and $\{s_n\}$ increases to h. For each n, we have just proved that $\int_E s_n\,d\mu = \int_E t_n\,d\mu + \int_E u_n\,d\mu$, and therefore $\int_E h\,d\mu = \int_E f\,d\mu + \int_E g\,d\mu$ by the Monotone Convergence Theorem (Theorem 5.25).

The next case is that $f \geq 0$, $g \leq 0$, and $h = f + g \geq 0$. Then $f = h + (-g)$ with $h \geq 0$ and $(-g) \geq 0$, so that $\int_E f\,d\mu = \int_E h\,d\mu + \int_E (-g)\,d\mu$. Hence $\int_E h\,d\mu = \int_E f\,d\mu + \int_E g\,d\mu$, provided the right side is defined.

For a general $h \geq 0$, we decompose E into the disjoint union of three sets, one where $f \geq 0$ and $g \geq 0$, one where $f \geq 0$ and $g < 0$, and one where $f < 0$ and $g \geq 0$. The additivity of the integral as a set function (Theorem 5.19), in combination with the cases that we have already proved, then gives the desired result. Finally for general h, we have only to write $h = h^+ - h^-$ and consider h^+ and h^- separately.

Corollary 5.27. Let E be a measurable set, and let $\{f_n\}$ be a sequence of measurable functions ≥ 0. Put $F(x) = \sum_{n=1}^{\infty} f_n(x)$. Then $\int_E F\,d\mu = \sum_{n=1}^{\infty} \int_E f_n\,d\mu$.

PROOF. Apply Proposition 5.26 to the n^{th} partial sum of the series, and then use the Monotone Convergence Theorem (Theorem 5.25).

The next corollary is given partly to illustrate a standard technique for passing from integration results about indicator functions to integration results about general functions. This technique is used again and again in measure theory.

Corollary 5.28. If $f \geq 0$ is a measurable function and if ν is the measure $\nu(E) = \int_E f\,d\mu$, then $\int_E g\,d\nu = \int_E gf\,d\mu$ for every measurable function g for which at least one side is defined.

REMARKS. The set function ν is a measure by Theorem 5.19. In the situation of this corollary, we shall write $\nu = f\ d\mu$.

PROOF. By Corollary 5.20 it is enough to prove that

$$\int_X g\,d\nu = \int_X gf\,d\mu. \tag{$*$}$$

For $g = I_E$, $(*)$ is true by hypothesis. Proposition 5.26 shows that $(*)$ extends to be valid for simple functions $g \geq 0$. For general $g \geq 0$, Proposition 5.11 produces an increasing sequence $\{s_n\}$ of simple functions ≥ 0 with pointwise limit g. Then $(*)$ for this g follows from the result for simple functions in combination with monotone convergence. For general g, write $g = g^+ - g^-$, apply $(*)$ for g^+ and g^-, and subtract the results using Proposition 5.26.

Theorem 5.29 (Fatou's Lemma). If E is a measurable set and if $\{f_n\}$ is a sequence of *nonnegative* measurable functions, then

$$\int_E \liminf_n f_n\,d\mu \leq \liminf_n \int_E f_n\,d\mu.$$

In particular, if $f(x) = \lim_n f_n(x)$ exists for all x, then

$$\int_E f\,d\mu \leq \liminf_n \int_E f_n\,d\mu.$$

REMARK. Fatou's Lemma applies to both examples that precede the Monotone Convergence Theorem (Theorem 5.25), and strict inequality holds in both cases.

PROOF. Set $g_n(x) = \inf_{k \geq n} f_k(x)$. Then $\lim_n g_n(x) = \liminf f_n(x)$, and the Monotone Convergence Theorem (Theorem 5.25) gives

$$\int_E \liminf_n f_n\,d\mu = \int_E \lim_n g_n\,d\mu = \lim_n \int_E g_n\,d\mu.$$

But $g_n(x) \leq f_n(x)$ pointwise, so that $\int_E g_n\,d\mu \leq \int_E f_n\,d\mu$ for all n. Thus

$$\lim \int_E g_n\,d\mu \leq \liminf \int_E f_n\,d\mu,$$

and the theorem follows.

Theorem 5.30 (Dominated Convergence Theorem). Let E be a measurable set, and suppose that $\{f_n\}$ is a sequence of measurable functions such that for some *integrable* g, $|f_n| \le g$ for all n. If $f = \lim f_n$ exists pointwise, then $\lim_n \int_E f_n \, d\mu$ exists, f is integrable on E, and

$$\int_E f \, d\mu = \lim_n \int_E f_n \, d\mu.$$

PROOF. The set on which g is infinite has measure 0, since g is integrable. If we redefine g, f_n, and f to be 0 on this set, we change no integrals and we affect the validity of neither the hypotheses nor the conclusion.

By Corollary 5.18, f is integrable on E, and so is f_n for every n. Applying Fatou's Lemma (Theorem 5.29) to $f_n + g \ge 0$, we obtain $\int_E (f + g) \, d\mu \le \liminf \int_E (f_n + g) \, d\mu$. Since g is integrable and everywhere finite, this inequality becomes

$$\int_E f \, d\mu \le \liminf \int_E f_n \, d\mu.$$

A second application of Fatou's Lemma, this time to $g - f_n \ge 0$, gives $\int_E (g - f) \, d\mu \le \liminf \int_E (g - f_n) \, d\mu$. Thus

$$-\int_E f \, d\mu \le \liminf \int_E (-f_n) \, d\mu$$

and

$$\int_E f \, d\mu \ge \limsup \int_E f_n \, d\mu.$$

Therefore $\lim \int_E f_n \, d\mu$ exists and has the value asserted.

Corollary 5.31. Let E be a set of finite measure, let $c \ge 0$ be in $\mathbb{R}$, and suppose that $\{f_n\}$ is a sequence of measurable functions such that $|f_n| \le c$ for all n. If $f = \lim f_n$ exists pointwise, then $\lim \int_E f_n \, d\mu$ exists, f is integrable on E, and

$$\int_E f \, d\mu = \lim_n \int_E f_n \, d\mu.$$

PROOF. This is the special case $g = c$ in Theorem 5.30.

5. Proof of the Extension Theorem

In this section we shall prove the Extension Theorem, Theorem 5.5. After the end of the proof, we shall fill in one further detail left from Section 1—to show that a measure on a σ-ring has a canonical extension to a measure on the smallest σ-algebra containing the given σ-ring.

Most of this section will concern the proof of the Extension Theorem in the case that X is measurable and $\nu(X)$ is finite. Thus, until further notice, let us assume that X is a nonempty set, $\mathcal{A}$ is an algebra of subsets of X, and ν is a nonnegative completely additive set function defined on $\mathcal{A}$ such that $\nu(X) < +\infty$.

In a way, the intuition for the proof is typical of that for many existence-uniqueness theorems in mathematics: to see how to prove existence, we assume existence and uniqueness outright, see what necessary conditions each of the assumptions puts on the object to be constructed, and then begin the proof.

With the present theorem in the case that $\nu(X)$ is finite, we shall assign to each subset E of X an upper bound $\mu^*(E)$ and a lower bound $\mu_*(E)$ for the value of the extended measure on the set E. If the existence half of the theorem is valid, we must have $\mu_*(E) \le \mu^*(E)$ for E in the smallest σ-algebra containing $\mathcal{A}$. In fact, we shall see that this inequality holds for *all* subsets E of X. On the other hand, if $\mu_*(E) < \mu^*(E)$ for some E in the σ-algebra of interest and if our upper and lower bounds are good estimates, we might expect that there is more than one way to define the extended measure on E, in contradiction to uniqueness. That thought suggests trying to prove that $\mu_*(E) = \mu^*(E)$ for the sets of interest. One way of doing so is to try to prove that the class of subsets for which this equality holds is a σ-algebra containing $\mathcal{A}$, and then the common value of μ_* and μ^* is the desired extension.

This procedure in fact works, and the only subtlety is in the definitions of $\mu_*(E)$ and $\mu^*(E)$. We give these definitions after one preliminary lemma that will make μ_* and μ^* well defined. For orientation, think of the setting as the unit interval $[0, 1]$, with Lebesgue measure to be extended from the elementary sets to the Borel sets. In this case the families $\mathcal{U}$ and $\mathcal{K}$ in the first lemma contain all the open sets and all the compact sets, respectively, and may be regarded as generalizations of these collections of sets.

Lemma 5.32. Let $\mathcal{U}$ be the class of all countable unions of sets in $\mathcal{A}$, and let $\mathcal{K}$ be the class of all countable intersections of sets in $\mathcal{A}$. Then μ^* and μ_* are consistently defined on $\mathcal{U}$ and $\mathcal{K}$, respectively, by letting

$$\mu^*(U) = \lim \nu(A_n) \qquad \text{and} \qquad \mu_*(K) = \lim \nu(C_n)$$

whenever $\{A_n\}$ is an increasing sequence of sets in $\mathcal{A}$ with union U and $\{C_n\}$ is a decreasing sequence of sets in $\mathcal{A}$ with intersection K. Moreover, μ^* and μ_* have the following properties:

(a) μ^* and μ_* agree with ν on sets of $\mathcal{A}$,
(b) $\mu^*(U) \le \mu^*(V)$ whenever U is in $\mathcal{U}$, V is in $\mathcal{U}$, and $U \subseteq V$,

(c) $\mu_*(K) \le \mu_*(L)$ whenever K is in $\mathcal{K}$, L is in $\mathcal{K}$, and $K \subseteq L$,

(d) $\lim \mu^*(U_n) = \mu^*(U)$ whenever $\{U_n\}$ is an increasing sequence of sets in $\mathcal{U}$ with union U.

PROOF. If $\{B_n\}$ is another increasing sequence in $\mathcal{A}$ with union U, then Proposition 5.2 and Theorem 1.13 give

$$\lim_m \nu(A_m) = \lim_m \big(\lim_n \nu(A_m \cap B_n)\big) = \lim_n \big(\lim_m \nu(A_m \cap B_n)\big) = \lim_n \nu(B_n).$$

Hence μ^* is consistently defined on $\mathcal{U}$. Similarly if $\{D_n\}$ decreases to K, then Corollary 5.3 and Theorem 1.13 give

$$\begin{aligned}\nu(X) - \lim_m \nu(C_m) &= \nu(X) - \lim_m \big(\lim_n \nu(C_m \cap D_n)\big)\\ &= \nu(X) - \lim_n \big(\lim_m \nu(C_m \cap D_n)\big) = \nu(X) - \lim_n \nu(D_n),\end{aligned}$$

and hence $\lim_m \nu(C_m) = \lim_n \nu(D_n)$. Thus μ_* is consistently defined on $\mathcal{K}$. The set functions μ^* and μ_* are defined on all of $\mathcal{U}$ and $\mathcal{K}$ because a set that is a countable union (or intersection) of sets in an algebra is a countable increasing union (or decreasing intersection).

Of the four properties, (a) is clear, and (b) and (c) follow from the inequalities

$$\mu^*(U) = \sup_{A \subseteq U,\, A \in \mathcal{A}} \nu(A) \le \sup_{A \subseteq V,\, A \in \mathcal{A}} \nu(A) = \mu^*(V)$$

and

$$\mu_*(K) = \inf_{A \supseteq K,\, A \in \mathcal{A}} \nu(A) \le \inf_{A \supseteq L,\, A \in \mathcal{A}} \nu(A) = \mu_*(L).$$

In (d), U is in $\mathcal{U}$, since the countable union of countable unions is again a countable union, and (b) shows that $\lim \mu^*(U_n) \le \mu^*(U)$. For each n, let $\{A_m^{(n)}\}$ be an increasing sequence of sets from $\mathcal{A}$ with union U_n. Arrange all the $A_m^{(n)}$ in a sequence, and let B_k denote the union of the first k members of the sequence. Then $\{B_k\}$ is an increasing sequence with union U. Let $\epsilon > 0$ be given, and choose M large enough so that $\mu^*(B_M) \ge \mu^*(U) - \epsilon$. Since the sets U_n increase, since B_M is a finite union of sets $A_m^{(n)}$, and since $A_m^{(n)} \subseteq U_n$, we must have $\mu^*(U_N) \ge \mu^*(B_M)$ for some N. But then

$$\lim \mu^*(U_n) \ge \mu^*(U_N) \ge \mu^*(B_M) \ge \mu^*(U) - \epsilon.$$

Since ϵ is arbitrary, $\lim \mu^*(U_n) \ge \mu^*(U)$.

For each subset E of X, we define

$$\mu^*(E) = \inf_{U \supseteq E,\, U \in \mathcal{U}} \mu^*(U) \qquad \text{and} \qquad \mu_*(E) = \sup_{K \subseteq E,\, K \in \mathcal{K}} \mu_*(K).$$

Conclusions (b) and (c) of Lemma 5.32 show that the new definitions of μ^* and μ_* are consistent with the old ones. The set functions μ^* and μ_* on arbitrary subsets E of X may be called the **outer measure** and the **inner measure** associated to ν.

Lemma 5.33. If A and B are subsets of X with $A \subseteq B$, then $\mu^*(A) \le \mu^*(B)$ and $\mu_*(A) \le \mu_*(B)$. In addition,

(a) if $E \subseteq \bigcup_{n=1}^\infty E_n$, then $\mu^*(E) \le \sum_{n=1}^\infty \mu^*(E_n)$,
(b) if F and G are disjoint, then $\mu_*(F) + \mu_*(G) \le \mu_*(F \cup G)$.

PROOF. Since $\mu^*(A)$ is an infimum over a larger class of sets than $\mu^*(B)$ is, we have $\mu^*(A) \le \mu^*(B)$. Similarly $\mu_*(A) \le \mu_*(B)$.

For (a), let $E \subseteq \bigcup_{n=1}^\infty E_n$. In the special case in which E_n is in $\mathcal{U}$ for all n, let $\{F_m^{(n)}\}$ be, for fixed n and varying m, an increasing sequence of sets in $\mathcal{A}$ with union E_n. For any N, we then have $\bigcup_{m=1}^\infty (F_m^{(1)} \cup \cdots \cup F_m^{(N)}) = E_1 \cup \cdots \cup E_N$. Hence

$$\begin{aligned}
\mu^*(E) \le \mu^*\Big(\bigcup_{n=1}^\infty E_n\Big) &= \lim_N \mu^*\Big(\bigcup_{n=1}^N E_n\Big) && \text{by Lemma 5.32d} \\
&= \lim_N \mu^*\Big(\bigcup_{m=1}^\infty (F_m^{(1)} \cup \cdots \cup F_m^{(N)})\Big) \\
&= \lim_N \lim_m \nu(F_m^{(1)} \cup \cdots \cup F_m^{(N)}) && \text{by definition of } \mu^* \text{ on } \mathcal{U} \\
&\le \lim_N \lim_m \sum_{n=1}^N \nu(F_m^{(n)}) && \text{by Proposition 5.1f} \\
&= \lim_N \sum_{n=1}^N \mu^*(E_n) = \sum_{n=1}^\infty \mu^*(E_n).
\end{aligned}$$

For general subsets E_n of X, choose U_n in $\mathcal{U}$ with $U_n \supseteq E_n$ and $\mu^*(U_n) \le \mu^*(E_n) + \epsilon/2^n$. Then $E \subseteq \bigcup_n U_n$, and the special case applied to the U_n shows that

$$\mu^*(E) \le \mu^*\Big(\bigcup_n U_n\Big) \le \sum_n \mu^*(U_n) \le \sum_n \mu^*(E_n) + \epsilon.$$

Hence $\mu^*(E) \le \sum_n \mu^*(E_n)$, and (a) is proved.

For (b), let F and G be disjoint. In the special case in which F and G are in $\mathcal{K}$, let $\{F_n\}$ and $\{G_n\}$ be decreasing sequences of sets in $\mathcal{A}$ with intersections F and G. Then

$$\begin{aligned}
\mu_*(F \cup G) &= \lim \nu(F_n \cup G_n) && \text{by definition of } \mu_* \text{ on } \mathcal{K} \\
&= \lim \big(\nu(F_n) + \nu(G_n) - \nu(F_n \cap G_n)\big) && \text{by Proposition 5.1b} \\
&= \mu_*(F) + \mu_*(G),
\end{aligned}$$

the last step holding by Corollary 5.3, since $F \cap G$ is empty. For general disjoint subsets F and G in X, choose K and L in $\mathcal{K}$ with $K \subseteq F$, $L \subseteq G$, $\mu_*(K) \geq \mu_*(F) - \epsilon$, and $\mu_*(L) \geq \mu_*(G) - \epsilon$. Then

$$\mu_*(F \cup G) \geq \mu_*(K \cup L) = \mu_*(K) + \mu_*(L) \geq \mu_*(F) + \mu_*(G) - 2\epsilon,$$

the middle step holding by the special case. Hence $\mu_*(F \cup G) \geq \mu_*(F) + \mu_*(G)$, and (b) is proved.

Lemma 5.34. For every subset E of X, $\mu_*(E) \leq \mu^*(E)$. Equality holds if E is in $\mathcal{U}$ or $\mathcal{K}$.

PROOF. The proof is in three steps.

First we prove that if U is in $\mathcal{U}$ and K is in $\mathcal{K}$, then $\mu^*(U) \leq \mu_*(U)$ and $\mu^*(K) \leq \mu_*(K)$. In fact, choose C in $\mathcal{A}$ with $C \subseteq U$ and $\mu^*(U) \leq \nu(C) + \epsilon$. Then $\mu^*(U) \leq \nu(C) + \epsilon \leq \mu_*(U) + \epsilon$ by Lemma 5.33 since $C \subseteq U$. Hence $\mu^*(U) \leq \mu_*(U)$. Similarly choose D in $\mathcal{A}$ with $D \supseteq K$ and $\mu_*(K) \geq \nu(D) - \epsilon$. Then $\mu_*(K) \geq \nu(D) - \epsilon \geq \mu^*(K) - \epsilon$, and hence $\mu_*(K) \geq \mu^*(K)$.

Second we prove that if K is in $\mathcal{K}$, then $\mu^*(K) = \mu_*(K)$. In fact, choose C in $\mathcal{A}$ with $C \supseteq K$ and $\nu(C) - \mu_*(K) \leq \epsilon$. Then $C - K$ is in $\mathcal{U}$, and

$$\begin{aligned}
\mu_*(K) \leq \nu(C) &\leq \mu^*(C - K) + \mu^*(K) && \text{by Lemma 5.33a} \\
&\leq \big(\mu_*(C - K) + \mu_*(K)\big) - \mu_*(K) + \mu^*(K) && \text{by the previous step} \\
&\leq \nu(C) - \mu_*(K) + \mu^*(K) && \text{by Lemma 5.33b} \\
&\leq \mu^*(K) + \epsilon && \text{by the choice of } C.
\end{aligned}$$

Combining this inequality with the previous step, we see that $\mu^*(K) = \mu_*(K)$.

Third we prove that $\mu_*(E) \leq \mu^*(E)$ for every E. In fact, find K in $\mathcal{K}$ and U in $\mathcal{U}$ with $K \subseteq E \subseteq U$, $\mu_*(K) \geq \mu_*(E) - \epsilon$, and $\mu^*(U) \leq \mu^*(E) + \epsilon$. Then $\mu_*(E) \leq \mu_*(K) + \epsilon = \mu^*(K) + \epsilon \leq \mu^*(U) + \epsilon \leq \mu^*(E) + 2\epsilon$, and the proof is complete.

Define a subset E of X to be **measurable** for purposes of this section if $\mu_*(E) = \mu^*(E)$, and let $\mathcal{B}$ be the class of measurable subsets of X. Lemma 5.34 shows that $\mathcal{U}$ and $\mathcal{K}$ are both contained in $\mathcal{B}$.

Lemma 5.35. If U is in $\mathcal{U}$ and K is in $\mathcal{K}$ with $K \subseteq U$, then

$$\mu^*(U - K) = \mu^*(U) - \mu_*(K).$$

If E is measurable, then for any $\epsilon > 0$, there are sets K in $\mathcal{K}$ and U in $\mathcal{U}$ with $K \subseteq E \subseteq U$ and

$$\mu^*(E - K) \leq \mu^*(U - K) \leq \epsilon.$$

PROOF. For the first conclusion, $U - K$ is in $\mathcal{U}$ and hence $\mu^*(U - K) = \mu_*(U - K) = \mu_*(U) - \mu_*(K) = \mu^*(U) - \mu_*(K)$ by Lemma 5.34, Lemma 5.33b, and Lemma 5.34 again.

For the second conclusion choose K in $\mathcal{K}$ and U in $\mathcal{U}$ with $K \subseteq E \subseteq U$, $\mu_*(K) + \frac{\epsilon}{2} \geq \mu_*(E)$, and $\mu^*(E) \geq \mu^*(U) - \frac{\epsilon}{2}$. Since $\mu_*(E) = \mu^*(E)$ by the assumed measurability, we see that $\mu_*(K) + \frac{\epsilon}{2} \geq \mu^*(U) - \frac{\epsilon}{2}$, hence that $\mu^*(U) - \mu_*(K) \leq \epsilon$. The result now follows from Lemma 5.33 and the first conclusion of the present lemma.

Lemma 5.36. The class $\mathcal{B}$ of measurable sets is a σ-algebra containing $\mathcal{A}$, and the restriction of μ^* to $\mathcal{B}$ is a measure.

PROOF. Certainly $\mathcal{B} \supseteq \mathcal{A}$. The rest of the proof is in three steps.

First we prove that the intersection of two measurable sets is measurable. In fact, let F and G be in $\mathcal{B}$, and use Lemma 5.35 to choose $K \subseteq F$ and $L \subseteq G$ with $\mu^*(F-K) \leq \epsilon$ and $\mu^*(G-L) \leq \epsilon$. Since $F \cap G \subseteq (F-K) \cup (K \cap L) \cup (G-L)$,

$$\begin{aligned}
&\mu^*(F \cap G) \\
&\qquad \leq \mu^*(F-K) + \mu^*(K \cap L) + \mu^*(G-L) && \text{by Lemma 5.33a} \\
&\qquad \leq \mu^*(K \cap L) + 2\epsilon && \text{by definition of } K \text{ and } L \\
&\qquad = \mu_*(K \cap L) + 2\epsilon && \text{by Lemma 5.34} \\
&\qquad \leq \mu_*(F \cap G) + 2\epsilon && \text{since } K \cap L \subseteq F \cap G.
\end{aligned}$$

Second we prove that the complement of a measurable set is measurable. Let E be measurable. By Lemma 5.35 choose K in $\mathcal{K}$ and U in $\mathcal{U}$ with $K \subseteq E \subseteq U$ and $\mu^*(U - K) \leq \epsilon$. Since $U^c \subseteq E^c \subseteq K^c$ and $K^c - U^c = U - K$, we have

$$\begin{aligned}
\mu^*(E^c) &\leq \mu^*(K^c - U^c) + \mu^*(U^c) && \text{by Lemma 5.33a} \\
&= \mu^*(U - K) + \mu_*(U^c) && \text{since } U^c \text{ is in } \mathcal{K} \\
&\leq \epsilon + \mu_*(E^c).
\end{aligned}$$

Thus the complement of a measurable set is measurable, and $\mathcal{B}$ is an algebra of sets.

Third we prove that the countable disjoint union of measurable sets is measurable, and μ^* is a measure on $\mathcal{B}$. In fact, let $\{E_n\}$ be a sequence of disjoint sets in $\mathcal{B}$. Application of Lemma 5.33a, Lemma 5.33b, and Lemma 5.34 gives

$$\begin{aligned}
\mu^*\Big(\bigcup_{n=1}^{\infty} E_n\Big) &\leq \sum_{n=1}^{\infty} \mu^*(E_n) = \sum_{n=1}^{\infty} \mu_*(E_n) = \lim_N \sum_{n=1}^{N} \mu_*(E_n) \\
&\leq \lim_N \mu_*\Big(\bigcup_{n=1}^{N} E_n\Big) \leq \mu_*\Big(\bigcup_{n=1}^{\infty} E_n\Big) \leq \mu^*\Big(\bigcup_{n=1}^{\infty} E_n\Big).
\end{aligned}$$

The end members of this chain of inequalities are equal, and thus equality must hold throughout: $\mu_*(\bigcup_n E_n) = \mu^*(\bigcup_n E_n) = \sum \mu^*(E_n)$. Consequently $\bigcup_n E_n$ is measurable, and μ^* is completely additive.

PROOF OF THEOREM 5.5 UNDER THE SPECIAL HYPOTHESES. We continue to assume that the given ring of subsets of X is an algebra and that $\nu(X)$ is finite. Define $\mathcal{B}$ to be the class of measurable sets in the previous construction. Then Lemma 5.36 shows that $\mathcal{B}$ is a σ-algebra containing $\mathcal{A}$. Hence $\mathcal{B}$ contains the smallest σ-algebra $\mathcal{C}$ containing $\mathcal{A}$. Lemma 5.36 shows also that the restriction of μ^* to $\mathcal{C}$ is a measure extending ν. This proves existence of the extension under the special hypotheses.

For uniqueness, suppose that μ' is an extension of ν to $\mathcal{C}$. Proposition 5.2 and Corollary 5.3 show that μ' has to agree with μ^* on $\mathcal{U}$ and with μ_* on $\mathcal{K}$. If $K \subseteq E \subseteq U$ with K in $\mathcal{K}$ and U in $\mathcal{U}$, then we have

$$\mu_*(K) = \mu'(K) \le \mu'(E) \le \mu'(U) = \mu^*(U).$$

Taking the supremum over K and the infimum over U gives $\mu_*(E) \le \mu'(E) \le \mu^*(E)$. Since E is in $\mathcal{B}$, $\mu_*(E) = \mu^*(E)$, and we see that $\mu'(E) = \mu^*(E)$. Thus μ' coincides with the restriction of μ^* to $\mathcal{C}$. This proves uniqueness of the extension under the special hypotheses.

Now we return to the general hypotheses of Theorem 5.5—that $\mathcal{R}$ is a ring of subsets of X, that ν is a nonnegative completely additive set function on $\mathcal{R}$, and that ν is σ-finite—and we shall complete the proof that ν extends uniquely to a measure on the smallest σ-ring $\mathcal{C}$ containing $\mathcal{R}$.

PROOF OF THEOREM 5.5 IN THE GENERAL CASE. If S is an element of $\mathcal{R}$ with $\nu(S)$ finite, define $S \cap \mathcal{R} = \{S \cap R \mid R \in \mathcal{R}\}$. Then $(S, S \cap \mathcal{R}, \nu\big|_{S\cap\mathcal{R}})$ is a set of data satisfying the special hypotheses of the Extension Theorem considered above. By the special case, if $\mathcal{C}_S$ denotes the smallest σ-algebra of subsets of S containing $S \cap \mathcal{R}$, then $\nu\big|_{S\cap\mathcal{R}}$ has a unique extension to a measure μ_S on $\mathcal{C}_S$. The measures μ_S have a certain consistency property because of the uniqueness: if $S' \subseteq S$, then $\mu_S\big|_{S'\cap\mathcal{R}} = \mu_{S'}$.

Now let $\{S_n\}$ be a sequence of sets in $\mathcal{R}$ with union S in $\mathcal{C}$ and with $\nu(S_n)$ finite for all n. Possibly replacing each set S_n by the difference of S_n and all previous S_k's, we may assume that the sequence is disjoint. We define μ_S on the σ-algebra $S \cap \mathcal{C}$ of subsets S by $\mu_S(E) = \sum_n \mu_{S_n}(E \cap S_n)$ for E in $S \cap \mathcal{C}$. Let us check that μ_S is unambiguously defined and is completely additive. If $\{T_m\}$ is another sequence of sets in $\mathcal{R}$ with union S and with $\nu(T_m)$ finite for all m, then the corresponding definition of a set function on $S \cap \mathcal{C}$ is $\mu'_S(E) = \sum_m \mu_{T_m}(E \cap T_m)$. The consistency property from the previous paragraph gives us $\mu_{S_n}(E \cap S_n \cap T_m) = \mu_{T_m}(E \cap S_n \cap T_m)$. Then Corollary 1.15 allows us to write

$$\begin{aligned}
\mu'_S(E) &= \sum_m \mu_{T_m}(E \cap T_m) = \sum_m \sum_n \mu_{T_m}(E \cap S_n \cap T_m) \\
&= \sum_m \sum_n \mu_{S_n}(E \cap S_n \cap T_m) = \sum_n \sum_m \mu_{S_n}(E \cap S_n \cap T_m) \\
&= \sum_n \mu_{S_n}(E \cap S_n) = \mu_S(E),
\end{aligned}$$

and we see that μ_S is unambiguously defined. To check that μ_S is completely additive, let $F_1, F_2, \dots$ be a disjoint sequence of sets in $S \cap \mathcal{C}$ with union F. Then the complete additivity of μ_{S_n}, in combination with Corollary 1.15, gives

$$\begin{aligned}
\mu_S(F) &= \sum_n \mu_{S_n}(F \cap S_n) = \sum_n \sum_m \mu_{S_n}(F_m \cap S_n) \\
&= \sum_m \sum_n \mu_{S_n}(F_m \cap S_n) = \sum_m \mu_S(F_m),
\end{aligned}$$

and thus μ_S is completely additive.

The measures μ_S are consistent on their common domains. To see the consistency, let us see that μ_S and μ_T agree on subsets of $S \cap T$. Let S be the countable disjoint union of sets S_n in $\mathcal{R}$, and let T be the countable disjoint union of sets T_m in $\mathcal{R}$. Then $S \cap T$ is the countable disjoint union of the sets $S_n \cap T_m$. If E is in $(S \cap T) \cap \mathcal{C}$, then Corollary 1.15 and the consistency property of the set functions μ_R for R in $\mathcal{R}$ yield

$$\begin{aligned}
\mu_S(E) &= \sum_n \mu_{S_n}(E \cap S_n) = \sum_n \mu_{S_n}(E \cap S_n \cap T) \\
&= \sum_n \sum_m \mu_{S_n}(E \cap S_n \cap T_m) = \sum_n \sum_m \mu_{S_n \cap T_m}(E \cap S_n \cap T_m) \\
&= \sum_m \sum_n \mu_{S_n \cap T_m}(E \cap S_n \cap T_m) = \sum_m \sum_n \mu_{T_m}(E \cap S_n \cap T_m) \\
&= \sum_m \mu_{T_m}(E \cap S \cap T_m) = \sum_m \mu_{T_m}(E \cap T_m) = \mu_T(E).
\end{aligned}$$

Hence the measures μ_S are consistent on their common domains.

If $\mathcal{M}$ denotes the set of subsets of X that are contained in a countable union of members of $\mathcal{R}$ on which ν is finite, then $\mathcal{M}$ is closed under countable unions and differences and is thus a σ-ring containing $\mathcal{R}$. It therefore contains $\mathcal{C}$, and we conclude that every member of $\mathcal{C}$ is contained in a countable union of members of $\mathcal{R}$ on which ν is finite. It follows that we can define μ on all of $\mathcal{C}$ as follows: if E

is in $\mathcal{C}$, then E is contained in some countable union S of members of $\mathcal{R}$ on which ν is finite, and we define $\mu(E) = \mu_S(E)$. We have seen that the measures μ_S are consistently defined, and hence $\mu(E)$ is well defined. If a countable disjoint union $E = \bigcup_{N=1}^\infty E_n$ of sets in $\mathcal{C}$ is given, then all the sets in question lie in a single S, and we then have $\mu(E) = \mu_S(E) = \sum_{n=1}^\infty \mu_S(E_n) = \sum_{n=1}^\infty \mu(E_n)$. In other words, μ is completely additive. This proves existence.

For uniqueness let E be given in $\mathcal{C}$, and suppose that S is a member of $\mathcal{C}$ containing E and equal to the countable disjoint union of sets S_n in $\mathcal{R}$ with $\nu(S_n)$ finite for all n. We have seen that the value of $\mu(E \cap S_n) = \mu_{S_n}(E \cap S_n)$ is determined by $\nu\big|_{S_n \cap \mathcal{R}}$, hence by ν on $\mathcal{R}$. By complete additivity of μ, $\mu(E)$ is determined by the values of $\mu(E \cap S_n)$ for all n. Therefore μ on $\mathcal{C}$ is determined by ν on $\mathcal{R}$. This proves uniqueness.

As was promised, we shall now fill in one further detail left from Section 1—to show that a measure on a σ-ring has a canonical extension to a measure on the smallest σ-algebra containing the given σ-ring.

Proposition 5.37. Let $\mathcal{R}$ be a σ-ring of subsets of a nonempty set X, let $\mathcal{R}_c$ be the set of complements in X of the members of $\mathcal{R}$, and let $\mathcal{A}$ be the smallest σ-algebra containing $\mathcal{R}$. Then either

(i) $\mathcal{R} = \mathcal{R}_c = \mathcal{A}$ or
(ii) $\mathcal{R} \cap \mathcal{R}_c = \varnothing$ and $\mathcal{A} = \mathcal{R} \cup \mathcal{R}_c$.

In the latter case any measure μ on $\mathcal{R}$ has a canonical extension to a measure μ_1 on $\mathcal{A}$ given by $\mu_1(E) = \sup\big\{\mu(F) \bigm| F \in \mathcal{R} \text{ and } F \subseteq E\big\}$ for E in $\mathcal{R}_c$. This canonical extension has the property that any other extension μ_2 satisfies $\mu_2 \geq \mu_1$.

PROOF. If X is in $\mathcal{R}$, then $\mathcal{R}$ is closed under complements, since $\mathcal{R}$ is closed under differences; hence $\mathcal{R} = \mathcal{R}_c = \mathcal{A}$. If X is not in $\mathcal{R}$, then $\mathcal{R} \cap \mathcal{R}_c = \varnothing$ because any set E in the intersection has E^c in the intersection and then also $X = E \cup E^c$ in the intersection. In this latter case it is plain that $\mathcal{A} \supseteq \mathcal{R} \cup \mathcal{R}_c$. Thus (ii) will be the only alternative to (i) if it is proved that $\mathcal{B} = \mathcal{R} \cup \mathcal{R}_c$ is a σ-algebra. Certainly $\mathcal{B}$ is closed under complements. To see that $\mathcal{B}$ is closed under countable unions, we may assume, because $\mathcal{R}$ is a σ-ring, that we are to check the union of countably many sets with at least one in $\mathcal{R}_c$. Thus let $\{E_n\}$ be a sequence of sets in $\mathcal{R}$, and let $\{F_n\}$ be a sequence of sets in $\mathcal{R}_c$. Then $E = \bigcup_{n=1}^\infty E_n$ is in $\mathcal{R}$ and $F = \bigcap_{n=1}^\infty F_n^c$ is in $\mathcal{R}$, since $\mathcal{R}$ is a σ-ring. The union of the sets E_n and F_n in question is $E \cup F^c = (F - E)^c$, is exhibited as the complement of the difference of two sets in $\mathcal{R}$, and is therefore in $\mathcal{R}_c$. Thus $\mathcal{A}$ is closed under countable unions and is a σ-algebra.

In the case of (ii), let us see that μ_1 is a measure on $\mathcal{A}$. If we are to check the measure of a disjoint sequence of sets in $\mathcal{A}$, there is no problem if all the sets are in $\mathcal{R}$, since $\mu_1\big|_{\mathcal{R}} = \mu$ is completely additive. There cannot be as many as two of the sets in $\mathcal{R}_c$ because no two sets F_1 and F_2 in $\mathcal{R}_c$ are disjoint; in fact, $F_1 \cap F_2 = (F_1^c \cup F_2^c)^c$ exhibits the intersection as in $\mathcal{R}_c$, and the empty set is not a member of $\mathcal{R}_c$. Thus we may assume that the disjoint sequence consists of a sequence $\{E_n\}$ of sets in $\mathcal{R}$ and a single set F in $\mathcal{R}_c$. If $E = \bigcup_{n=1}^\infty E_n$, then $\mu_1(E) = \mu(E) = \sum_{n=1}^\infty \mu(E_n) = \sum_{n=1}^\infty \mu_1(E_n)$. So it is enough to see that $\mu_1(E \cup F) = \mu(E) + \mu_1(F)$. If E' is a subset of F that is in $\mathcal{R}$, then $\mu_1(E \cup F) \geq \mu(E \cup E') = \mu(E) + \mu(E')$. Taking the supremum over all such E' shows that $\mu_1(E \cup F) \geq \mu(E) + \mu_1(F)$. For the reverse inequality let S be a member of $\mathcal{R}$ contained in $E \cup F$. Then the sets $E \cap S$ and $F \cap S = S - F^c$ are in $\mathcal{R}$, and thus $\mu(S) = \mu(E \cap S) + \mu(F \cap S) \leq \mu(E) + \mu_1(F)$. Taking the supremum over S gives $\mu_1(E \cup F) \leq \mu(E) + \mu_1(F)$. Thus μ_1 is completely additive.

If μ_2 is any other extension, any set F in $\mathcal{R}^c$ has $\mu_2(F) \geq \mu_2(E) = \mu(E)$ for all subsets E of F that are in $\mathcal{R}$. Taking the supremum over E gives $\mu_2(F) \geq \mu_1(F)$, and thus $\mu_2 \geq \mu_1$ as set functions on $\mathcal{A}$.

6. Completion of a Measure Space

If $(X, \mathcal{A}, \mu)$ is a measure space, we define the **completion** of this space to be the measure space $(X, \overline{\mathcal{A}}, \overline{\mu})$ defined by

$$\overline{\mathcal{A}} = \left\{ E \,\Delta\, Z \;\middle|\; \begin{array}{l} E \text{ is in } \mathcal{A} \text{ and } Z \subseteq Z' \text{ for} \\ \text{some } Z' \in \mathcal{A} \text{ with } \mu(Z') = 0 \end{array} \right\},$$

$$\overline{\mu}(E \,\Delta\, Z) = \mu(E).$$

It is necessary to verify that the result is in fact a measure space, and we shall carry out this step in the proposition below. In the case of Lebesgue measure m on the line, when initially defined on the σ-algebra $\mathcal{A}$ of Borel sets, the sets in σ-algebra $\overline{\mathcal{A}}$ are said to be **Lebesgue measurable**.

Proposition 5.38. If $(X, \mathcal{A}, \mu)$ is a measure space, then the completion $(X, \overline{\mathcal{A}}, \overline{\mu})$ is a measure space. Specifically

(a) $\overline{\mathcal{A}}$ is a σ-algebra containing $\mathcal{A}$,
(b) the set function $\overline{\mu}$ is unambiguously defined on $\overline{\mathcal{A}}$, i.e., if $E_1 \,\Delta\, Z_1 = E_2 \,\Delta\, Z_2$ as above, then $\mu(E_1) = \mu(E_2)$,
(c) $\overline{\mu}$ is a measure on $\overline{\mathcal{A}}$, and $\overline{\mu}(E) = \mu(E)$ for all sets E in $\mathcal{A}$.

In addition,

(d) if $\widetilde{\mu}$ is any measure on $\overline{\mathcal{A}}$ such that $\widetilde{\mu}(E) = \mu(E)$ for all E in $\mathcal{A}$, then $\widetilde{\mu} = \overline{\mu}$ on $\overline{\mathcal{A}}$,

(e) if $\mu(X) < +\infty$ and if for $E \subseteq X$, $\mu_*(E)$ and $\mu^*(E)$ are defined by

$$\mu_*(E) = \sup_{A \subseteq E,\ A \in \mathcal{A}} \mu(A) \qquad \text{and} \qquad \mu^*(E) = \inf_{A \supseteq E,\ A \in \mathcal{A}} \mu(A),$$

then E is in $\overline{\mathcal{A}}$ if and only if $\mu_*(E) = \mu^*(E)$.

PROOF. For (a), certainly $\mathcal{A} \subseteq \overline{\mathcal{A}}$ because we can use $Z = Z' = \varnothing$ in the definition of $\overline{\mathcal{A}}$. Since $(E \,\Delta\, Z)^c = (E \,\Delta\, Z) \,\Delta\, X = (E \,\Delta\, X) \,\Delta\, Z = E^c \,\Delta\, Z$, $\overline{\mathcal{A}}$ is closed under complements.

To prove closure under countable unions, let us first prove that

$$\overline{\mathcal{A}} = \left\{ E \cup Z \;\middle|\; \begin{array}{l} E \text{ is in } \mathcal{A} \text{ and } Z \subseteq Z' \text{ for} \\ \text{some } Z' \in \mathcal{A} \text{ with } \mu(Z') = 0 \end{array} \right\}. \tag{$*$}$$

Thus let $E \cup Z$ be given, with $Z \subseteq Z'$. Then $E \cup Z = E \,\Delta\, (Z \,\Delta\, (E \cap Z))$ with $Z \,\Delta\, (E \cap Z) \subseteq Z'$. So $E \cup Z$ is in $\overline{\mathcal{A}}$. Conversely if $E \,\Delta\, Z$ is in $\overline{\mathcal{A}}$, we can write $E \,\Delta\, Z = (E - Z') \cup ((E \cap Z') - Z) \cup (Z - E))$ with $((E \cap Z') - Z) \cup (Z - E)) \subseteq Z'$, and then we see that $E \,\Delta\, Z$ is of the form $E'' \cup Z''$ with E'' in $\mathcal{A}$ and $Z'' \subseteq Z'$.

Returning to the proof of closure under countable unions, let $E_n \cup Z_n$ be given in $\overline{\mathcal{A}}$ with $Z_n \subseteq Z'_n$ and $\mu(Z'_n) = 0$. Then $\bigcup_n (E_n \cup Z_n) = \big(\bigcup_n E_n\big) \cup \big(\bigcup_n Z_n\big)$ with $\bigcup_n Z_n \subseteq \bigcup_n Z'_n$ and $\mu\big(\bigcup_n Z'_n\big) = 0$. In view of $(*)$, $\overline{\mathcal{A}}$ is therefore closed under countable unions.

For (b), we take as given that $E_1 \,\Delta\, Z_1 = E_2 \,\Delta\, Z_2$ with $Z_1 \subseteq Z'_1$, $Z_2 \subseteq Z'_2$, and $\mu(Z'_1) = \mu(Z'_2) = 0$. Then $(E_1 \,\Delta\, E_2) \,\Delta\, (Z_1 \,\Delta\, Z_2) = \varnothing$ and hence $E_1 \,\Delta\, E_2 = Z_1 \,\Delta\, Z_2 \subseteq Z'_1 \cup Z'_2$. Therefore $\mu(E_1 - E_2) \le \mu(E_1 \,\Delta\, E_2) \le \mu(Z'_1 \cup Z'_2) = 0$ and similarly $\mu(E_2 - E_1) = 0$. It follows that $\mu(E_1) = \mu(E_1 - E_2) + \mu(E_1 \cap E_2) = \mu(E_1 \cap E_2) = \mu(E_2 - E_1) + \mu(E_1 \cap E_2) = \mu(E_2)$, and $\overline{\mu}$ is unambiguously defined.

For (c), we see from $(*)$ that $\overline{\mu}$ can be defined equivalently by $\overline{\mu}(E \cup Z) = \mu(E)$ if $Z \subseteq Z'$ and $\mu(Z') = 0$. If a disjoint sequence $E_n \cup Z_n$ is given, then we find that $\overline{\mu}\big(\bigcup_n (E_n \cup Z_n)\big) = \overline{\mu}\big(\big(\bigcup_n E_n\big) \cup \big(\bigcup_n Z_n\big)\big) = \mu\big(\bigcup_n E_n\big) = \sum \mu(E_n) = \sum \overline{\mu}(E_n \cup Z_n)$, and complete additivity is proved. Taking $Z = \varnothing$ in the definition $\overline{\mu}(E \cup Z) = \mu(E)$, we obtain $\overline{\mu}(E) = \mu(E)$ for E in $\mathcal{A}$.

For (d), we use $(*)$ as the description of the sets in $\overline{\mathcal{A}}$. Let $E \cup Z$ be in $\overline{\mathcal{A}}$ with E in $\mathcal{A}$, $Z \subseteq Z'$, and Z' in $\mathcal{A}$ with $\mu(Z') = 0$. Then Proposition 5.1f gives $\widetilde{\mu}(E \cap Z) \le \widetilde{\mu}(Z) \le \widetilde{\mu}(Z') = \mu(Z') = 0$, so that $\widetilde{\mu}(E \cap Z) = \widetilde{\mu}(Z) = 0$. Meanwhile, Proposition 5.1b gives $\widetilde{\mu}(E \cup Z) + \widetilde{\mu}(E \cap Z) = \widetilde{\mu}(E) + \widetilde{\mu}(Z)$. Hence $\widetilde{\mu}(E \cup Z) = \widetilde{\mu}(E) = \mu(E)$.

For (e), it is immediate that $\mu_*(E) \le \mu^*(E)$ for every subset E of X. Let $E = C \cup Z$ be in $\overline{\mathcal{A}}$ with C in $\mathcal{A}$, $Z \subseteq Z'$, and Z' in $\mathcal{A}$ with $\mu(Z') = 0$. Then $\mu(C) \le \mu_*(E) \le \mu^*(E) \le \mu(C \cup Z') \le \mu(C) + \mu(Z') = \mu(C)$. Since the expressions at the ends are equal, we must have equality throughout, and therefore $\mu_*(E) = \mu^*(E)$.

In the converse direction let $\mu_*(E) = \mu^*(E)$. We can find a sequence of sets $A_n \in \mathcal{A}$ contained in E with $\lim \mu(A_n) = \mu_*(E)$, and we may assume without loss of generality that $\{A_n\}$ is an increasing sequence. Similarly we can find a decreasing sequence of sets $B_n \in \mathcal{A}$ containing E with $\lim \mu(B_n) = \mu^*(E)$. Let $A = \bigcup_n A_n$ and $B = \bigcap_n B_n$. When combined with the equality $\mu_*(E) = \mu^*(E)$, Proposition 5.2 and Corollary 5.3 show that $\mu(A) = \mu_*(E) = \mu^*(E) = \mu(B)$. Since $A \subseteq E \subseteq B$, we have $\mu(B-A) = \mu(B) - \mu(A) = 0$ and $E = A \cup (E-A)$ with $E - A \subseteq B - A$ and $\mu(B - A) = 0$. By $(*)$, E is in $\overline{\mathcal{A}}$.

A variant of Proposition 5.38e and its proof identifies the σ-algebra on which the extended measure is constructed in the proof of the Extension Theorem (Theorem 5.5) in the special case we considered. In the special case of the Extension Theorem, the given ring of sets is an algebra $\mathcal{A}$, and $\nu(X)$ is finite. The set function ν gets extended to a measure μ on a σ-algebra $\mathcal{B}$ that contains the smallest σ-algebra $\mathcal{C}$ containing $\mathcal{A}$. The sets of $\mathcal{B}$ are those for which $\mu_*(E) = \mu^*(E)$, where

$$\mu^*(E) = \inf_{U \supseteq E,\, U \in \mathcal{U}} \mu^*(U) \qquad \text{and} \qquad \mu_*(E) = \sup_{K \subseteq E,\, K \in \mathcal{K}} \mu_*(K),$$

$\mathcal{K}$ and $\mathcal{U}$ having been defined in terms of countable intersections and countable unions, respectively, from $\mathcal{A}$. The variant of Proposition 5.38e is that a subset E of X has $\mu_*(E) = \mu^*(E)$ if and only if E is of the form $C \cup Z$ with C in $\mathcal{C}$, $Z \subseteq Z'$, and Z' in $\mathcal{C}$ with $\mu(Z') = 0$. In other words, $(X, \mathcal{B}, \mu)$ is the completion of $(X, \mathcal{C}, \mu)$.

The proof is modeled on the proof of Proposition 5.38e. If $E = C \cup Z$ is a set in $\overline{\mathcal{C}}$ with C in $\mathcal{C}$, $Z \subseteq Z'$, and Z' in $\mathcal{C}$ with $\mu(Z') = 0$, then $\mu(C) \le \mu_*(E) \le \mu^*(E) \le \mu(C \cup Z') \le \mu(C) + \mu(Z') = \mu(C)$. We conclude that $\mu_*(E) = \mu^*(E)$.

In the converse direction let $\mu_*(E) = \mu^*(E)$. We can find an increasing sequence of sets $A_n \in \mathcal{K} \subseteq \mathcal{C}$ contained in E with $\lim \mu(A_n) = \mu_*(E)$, and we can find a decreasing sequence of sets $B_n \in \mathcal{U} \subseteq \mathcal{C}$ containing E with $\lim \mu(B_n) = \mu^*(E)$. Let $A = \bigcup_n A_n$ and $B = \bigcap_n B_n$. Arguing as in the proof of Proposition 5.38e, we have $\mu(A) = \mu_*(E) = \mu^*(E) = \mu(B)$, $\mu(B - A) = \mu(B) - \mu(A) = 0$, and $E = A \cup (E - A)$ with $E - A \subseteq B - A$ and $\mu(B - A) = 0$. Thus $E = C \cup Z$ with $C = A$ and $Z = E - A$.

This calculation has the following interesting consequence.

Proposition 5.39. In $\mathbb{R}^1$, the Lebesgue measurable sets of measure 0 are exactly the subsets E of $\mathbb{R}^1$ with the following property: for any $\epsilon > 0$, the set E can be covered by countably many intervals of total length less than ϵ.

PROOF. Within a bounded interval $[a, b]$, the above remarks apply and show that the Lebesgue measurable sets of measure 0 are the sets E with $\mu^*(E) = 0$, where $\mu^*(E) = \inf_{U \supseteq E,\, U \in \mathcal{U}} \mu^*(U)$. The sets U defining $\mu^*(E)$ are countable unions of intervals, and the proposition follows for subsets of any bounded interval $[a, b]$.

For general sets E in $\mathbb{R}^1$, if the covering condition holds, then Proposition 5.1g shows that E has Lebesgue measure 0. Conversely if E is Lebesgue measurable of measure 0, then $E \cap [-N, N]$ is a bounded set of measure 0 and can be covered by countably many intervals of arbitrarily small total length. Let us arrange that the total length is $< 2^{-N}\epsilon$. Taking the union of these sets of intervals as N varies, we obtain a cover of E by countably many intervals of total length less than ϵ.

7. Fubini's Theorem for the Lebesgue Integral

Fubini's Theorem for the Lebesgue integral concerns the interchange of order of integration of functions of two variables, just as with the Riemann integral in Section III.9. In the case of Euclidean space $\mathbb{R}^n$, we could have constructed Lebesgue measure in each dimension by a procedure similar to the one we used for $\mathbb{R}^1$. Then Fubini's Theorem relates integration of a function of $m+n$ variables over a set by either integrating in all variables at once or integrating in the first m variables first or integrating in the last n variables first. In the context of more general measure spaces, we need to develop the notion of the product of two measure spaces. This corresponds to knowing $\mathbb{R}^m$ and $\mathbb{R}^n$ with their Lebesgue measures and to constructing $\mathbb{R}^{m+n}$ with its Lebesgue measure.

In the theorem as we shall state it, we are given two measures spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$, and we assume that both μ and ν are σ-finite. We shall construct a product measure space$(X \times Y, \mathcal{A} \times \mathcal{B}, \mu \times \nu)$, and the formula in question will be

$$\begin{aligned}\int_{X \times Y} f\, d(\mu \times \nu) &\stackrel{?}{=} \int_X \Big[\int_Y f(x, y)\, d\nu(y) \Big] d\mu(x) \\ &\stackrel{?}{=} \int_Y \Big[\int_X f(x, y)\, d\mu(x) \Big] d\nu(y).\end{aligned}$$

This formula will be valid for $f \geq 0$ measurable with respect to $\mathcal{A} \times \mathcal{B}$.

The technique of proof will be the standard one indicated in connection with proving Corollary 5.28. We start with indicator functions, extend the result to simple functions by linearity, and pass to the limit by the Monotone Convergence

Theorem (Theorem 5.25). It is then apparent that the difficult step is the case that f is an indicator function. In fact, it is not even clear in this special case that the inside integral $\int_Y I_E(x, y)\,d\nu(y)$ is a measurable function of X, and this is the step that requires some work.

We begin by describing $\mathcal{A}\times\mathcal{B}$, the σ-algebra of measurable sets for the product $X \times Y$. Recall from Section A1 of the appendix that $X \times Y$ is defined as a set of ordered pairs. If $A \subseteq X$ and $B \subseteq Y$, then the set of ordered pairs that constitute $A \times B$ is a subset of $X \times Y$, and we call $A \times B$ a **rectangle**[6] in $X \times Y$. The sets A and B are called the **sides** of the rectangle.

Proposition 5.40. If $\mathcal{A}$ and $\mathcal{B}$ are algebras of subsets of nonempty sets X and Y, then the class $\mathcal{C}$ of all finite disjoint unions of rectangles $A \times B$ with A in $\mathcal{A}$ and B in $\mathcal{B}$ is an algebra of sets in $X \times Y$. In particular, a finite union of rectangles is a finite disjoint union.

PROOF. The intersection of the rectangles $R_1 = A_1 \times B_1$ and $R_2 = A_2 \times B_2$ is the rectangle $R = (A_1 \cap A_2) \times (B_1 \cap B_2)$ because both $R_1 \cap R_2$ and R coincide with the set $\big\{(x, y) \in X \times Y \bigm| x \in A_1,\ x \in A_2,\ y \in B_1,\ y \in B_2\big\}$. Therefore

$$\Big(\bigcup_{i=1}^{m}(A_i \times B_i)\Big) \cap \Big(\bigcup_{j=1}^{n}(C_j \times D_j)\Big) = \bigcup_{i,j}\ \big\{(A_i \cap C_j) \times (B_i \cap D_j)\big\},$$

and the right side is a disjoint union if both $\bigcup_i (A_i \times B_i)$ and $\bigcup_j (C_j \times D_j)$ are disjoint unions. Moreover, the right side is in $\mathcal{C}$ if both unions on the left are in $\mathcal{C}$. Therefore $\mathcal{C}$ is closed under finite intersections.

Certainly $\varnothing$ and $X \times Y$ are in $\mathcal{C}$. The identity

$$(X \times Y) - (A \times B) = \big((X - A) \times B\big) \cup \big(X \times (Y - B)\big)$$

exhibits the complement of a rectangle as a disjoint union of rectangles. Since the complement of a disjoint union is the intersection of the complements, $\mathcal{C}$ is closed under complementation. Thus $\mathcal{C}$ is an algebra of sets, and the proof is complete.

If $\mathcal{A}$ and $\mathcal{B}$ are σ-algebras in X and Y, then we denote the smallest σ-algebra containing the algebra $\mathcal{C}$ of the above proposition by $\mathcal{A} \times \mathcal{B}$. The set $X \times Y$, together with the σ-algebra $\mathcal{A} \times \mathcal{B}$, is called a product space. The **measurable sets** of $X \times Y$ are the sets of $\mathcal{A} \times \mathcal{B}$.

[6]The word "rectangle" was used with a different meaning in Chapter III, but there will be no possibility of confusion for now. Starting in Chapter VI, both kinds of rectangles will be in play; the ones in Chapter III can then be called "geometric rectangles" and the present ones can be called "abstract rectangles."

Let E be any set in $X \times Y$. The **section** E_x of E determined by x in X is defined by

$$E_x = \{y \mid (x, y) \text{ is in } E\}.$$

Similarly the section E^y determined by y in Y is

$$E^y = \{x \mid (x, y) \text{ is in } E\}.$$

The section E_x is a subset of Y, and the section E^y is a subset of X.

Lemma 5.41. Let $\{E_\alpha\}$ be a class of subsets of $X \times Y$, and let x be a point of X. Then

(a) $\big(\bigcup_\alpha E_\alpha\big)_x = \bigcup_\alpha (E_\alpha)_x$,
(b) $\big(\bigcap_\alpha E_\alpha\big)_x = \bigcap_\alpha (E_\alpha)_x$,
(c) $(E_\alpha - E_\beta)_x = (E_\alpha)_x - (E_\beta)_x$ and, in particular, $(E^c_\beta)_x = Y - (E_\beta)_x$.

PROOF. These facts are special cases of the identities at the end of Section A1 of the appendix for inverse images of functions. In this case the function in question is given by $f(y) = (x, y)$.

Proposition 5.42. Let $\mathcal{A}$ and $\mathcal{B}$ be σ-algebras in X and Y, and let E be a measurable set in $X \times Y$. Then every section E_x is a measurable set in Y, and every section E^y is a measurable set in X.

PROOF. We prove the result for sections E_x, the proof for E^y being completely analogous. Let $\mathcal{E}$ be the class of all subsets E of $X \times Y$ all of whose sections E_x are in $\mathcal{B}$. Then $\mathcal{E}$ contains all rectangles with measurable sides, since a section of a rectangle is either the empty set or one of the sides. By Lemma 5.41a, $\mathcal{E}$ is closed under finite unions. Hence $\mathcal{E}$ contains the algebra $\mathcal{C}$ of finite disjoint unions of rectangles with measurable sides. By parts (a) and (c) of Lemma 5.41, $\mathcal{E}$ is closed under countable unions and complements. It is therefore a σ-algebra containing $\mathcal{C}$ and thus contains $\mathcal{A} \times \mathcal{B}$.

A corollary of Proposition 5.42 is that a rectangle in $X \times Y$ is measurable if and only if its sides are measurable. The sufficiency follows from the fact that a rectangle with measurable sides is in $\mathcal{C}$, and the necessity follows from the proposition.

From now on, we shall adhere to the convention that a **rectangle** is always assumed to be measurable.

We turn to the implementation of the sketch of proof of Fubini's Theorem given earlier in this section. The basic question will be the equality of the iterated integrals in either order when the integrand is an indicator function. If E is a measurable set in $X \times Y$, then we know from Proposition 5.42 that E_x is a measurable subset of Y. In order to form the iterated integral

$$\int_X \Big[\int_Y I_E(x, y)\, d\nu(y) \Big]\, d\mu(x),$$

we compute the inside integral as $\nu(E_x)$, and we have to be able to form the outside integral, which is $\int_X \nu(E_x)\, d\mu(x)$. That is, we need to know that $\nu(E_x)$ is a measurable function on X. For the iterated integral in the other order, we need to know that $\mu(E^y)$ is measurable on Y.

The proof of this measurability is the hard step, since the class of sets E for which $\nu(E_x)$ and $\mu(E^y)$ are both measurable does not appear to be necessarily a σ-algebra, even when μ and ν are finite measures. To deal with this difficulty, we introduce the following terminology: a class of sets is called a **monotone class** if it is closed under countable increasing unions and countable decreasing intersections. It is readily verified that the class of all subsets of a set is a monotone class and that the intersection of any nonempty family of monotone classes is a monotone class; hence there is a smallest monotone class containing any given class of sets.

The proof of the lemma below introduces the notation $\uparrow$ and $\downarrow$ to denote increasing countable union and decreasing countable intersection, respectively.

Lemma 5.43 (Monotone Class Lemma). The smallest monotone class $\mathcal{M}$ containing an algebra $\mathcal{A}$ of sets is identical to the smallest σ-algebra $\widetilde{\mathcal{A}}$ containing $\mathcal{A}$.

PROOF. We have $\mathcal{M} \subseteq \widetilde{\mathcal{A}}$ because $\widetilde{\mathcal{A}}$ is a monotone class containing $\mathcal{A}$. To prove the reverse inclusion, it is sufficient to show that $\mathcal{M}$ is closed under the operations of finite union and complementation, since a countable union can be written as the increasing countable union of finite unions. The proof is in three steps.

First we prove that if A is in $\mathcal{A}$ and M is in $\mathcal{M}$, then $A \cup M$ and $A \cap M$ are in $\mathcal{M}$. For fixed A in $\mathcal{A}$, let $\mathcal{U}_A$ be the class of all sets M in $\mathcal{M}$ such that $A \cup M$ and $A \cap M$ are in $\mathcal{M}$. Then $\mathcal{U}_A \supseteq \mathcal{A}$. If we show that $\mathcal{U}_A$ is a monotone class, then it will follow that $\mathcal{U}_A \supseteq \mathcal{M}$. For this purpose let

$$U_n \uparrow U \quad \text{and} \quad V_n \downarrow V \quad \text{with} \quad U_n \text{ and } V_n \text{ in } \mathcal{U}_A.$$

By definition of $\mathcal{U}_A$, the sets $U_n \cup A$, $U_n \cap A$, $V_n \cup A$, and $V_n \cap A$ are in $\mathcal{M}$. But

$$\begin{aligned} U_n \cup A \uparrow U \cup A \quad &\text{and} \quad U_n \cap A \uparrow U \cap A, \\ V_n \cup A \downarrow V \cup A \quad &\text{and} \quad V_n \cap A \downarrow V \cap A. \end{aligned}$$

Therefore U and V are in $\mathcal{U}_A$, and $\mathcal{U}_A$ is a monotone class.

Second we prove that $\mathcal{M}$ is closed under finite unions. For fixed N in $\mathcal{M}$, let $\mathcal{U}_N$ be the class of all sets M in $\mathcal{M}$ such that $N \cup M$ and $N \cap M$ are in $\mathcal{M}$. Then $\mathcal{U}_N \supseteq \mathcal{A}$ by the previous step. The same argument as in that step shows that $\mathcal{U}_N$ is a monotone class, and hence $\mathcal{U}_N = \mathcal{M}$.

Third we prove that $\mathcal{M}$ is closed under complements. Let $\mathcal{N}$ be the class of all sets in $\mathcal{M}$ whose complements are in $\mathcal{M}$. Then $\mathcal{N} \supseteq \mathcal{A}$, and it is enough to show that $\mathcal{N}$ is a monotone class. If

$$C_n \uparrow C \quad \text{and} \quad D_n \downarrow D \quad \text{with} \quad C_n \text{ and } D_n \text{ in } \mathcal{N},$$

then C and D are in $\mathcal{M}$ since C_n and D_n are in $\mathcal{M}$. Now

$$C_n^c \downarrow C^c \quad \text{and} \quad D_n^c \uparrow D^c,$$

and by definition of $\mathcal{N}$, C_n^c and D_n^c are in $\mathcal{M}$. Therefore C^c and D^c are in $\mathcal{M}$, and C and D must be in $\mathcal{N}$. That is, $\mathcal{N}$ is a monotone class.

Lemma 5.44. If $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are σ-finite measure spaces, then $\nu(E_x)$ and $\mu(E^y)$ are measurable functions for every E in $\mathcal{A} \times \mathcal{B}$.

PROOF IF $\mu(X) < +\infty$ AND $\nu(Y) < +\infty$. Let $\mathcal{M}$ be the class of all sets E in $\mathcal{A} \times \mathcal{B}$ for which $\nu(E_x)$ and $\mu(E^y)$ are measurable. We shall show that $\mathcal{M}$ is a monotone class containing the algebra $\mathcal{C}$ of finite disjoint unions of rectangles. If $R = A \times B$ is a rectangle, then

$$\nu(R_x) = \nu(B) I_A \qquad \text{and} \qquad \mu(R^y) = \mu(A) I_B,$$

and so R is in $\mathcal{M}$. If E and F are disjoint sets in $\mathcal{M}$, then

$$\nu((E \cup F)_x) = \nu(E_x \cup F_x) = \nu(E_x) + \nu(F_x)$$

for each x, and similarly for μ for each y. By Proposition 5.7, $\nu((E \cup F)_x)$ and $\mu((E \cup F)^y)$ are measurable. Hence $E \cup F$ is in $\mathcal{M}$, and $\mathcal{M}$ contains $\mathcal{C}$. If $\{E_n\}$ and $\{F_n\}$ are increasing and decreasing sequences of sets in $\mathcal{M}$, then the finiteness and complete additivity of ν imply that

$$\nu\Big(\Big(\bigcup_n E_n\Big)_x\Big) = \nu\Big(\bigcup_n (E_n)_x\Big) = \lim \nu((E_n)_x)$$

and

$$\nu\Big(\Big(\bigcap_n F_n\Big)_x\Big) = \nu\Big(\bigcap_n (F_n)_x\Big) = \lim \nu((F_n)_x),$$

and similarly for μ. Since the limit of measurable functions is measurable (Corollary 5.10), we conclude that $\mathcal{M}$ is a monotone class. Therefore $\mathcal{M}$ contains $\mathcal{A} \times \mathcal{B}$ by the Monotone Class Lemma (Lemma 5.43).

PROOF FOR σ-FINITE μ AND ν. Write $X = \bigcup_{m=1}^\infty X_m$ and $Y = \bigcup_{n=1}^\infty Y_n$ disjointly, with $\mu(X_m) < +\infty$ and $\nu(Y_n) < +\infty$ for all m and n. Define $\mathcal{A}_m$ and $\mathcal{B}_n$ by

$$\mathcal{A}_m = \big\{A \cap X_m \;\big|\; A \text{ is in } \mathcal{A}\big\} \quad \text{and} \quad \mathcal{B}_n = \big\{B \cap Y_n \;\big|\; B \text{ is in } \mathcal{B}\big\},$$

and define μ_m and ν_n on $\mathcal{A}_m$ and $\mathcal{B}_n$ by restriction from μ and ν. Then the triples $(X_m, \mathcal{A}_m, \mu_m)$ and $(Y_n, \mathcal{B}_n, \nu_n)$ are finite measure spaces, and the previous case applies. If E is in $\mathcal{A} \times \mathcal{B}$, then $E_{mn} = E \cap (X_m \times Y_n)$ is in $\mathcal{A}_m \times \mathcal{B}_n$, and so $\nu((E_{mn})_x)$ and $\mu((E_{mn})^y)$ are measurable with respect to $\mathcal{A}_m$ and $\mathcal{B}_n$, hence with respect to $\mathcal{A}$ and $\mathcal{B}$. Thus

$$\nu(E_x) = \sum_{m,n} \nu((E_{mn})_x) \qquad \text{and} \qquad \mu(E^y) = \sum_{m,n} \mu((E_{mn})^y)$$

exhibit $\nu(E_x)$ and $\mu(E^y)$ as countable sums of nonnegative measurable functions. They are therefore measurable.

The next proposition simultaneously constructs the product measure and establishes Fubini's Theorem for indicator functions.

Proposition 5.45. Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. Then there exists a unique measure $\mu \times \nu$ on $\mathcal{A} \times \mathcal{B}$ such that

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$$

for every rectangle $A \times B$. The measure $\mu \times \nu$ is σ-finite, and

$$(\mu \times \nu)(E) = \int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y)$$

for every set E in $\mathcal{A} \times \mathcal{B}$.

PROOF. In view of the measurability of $\nu(E_x)$ given in Lemma 5.44, we can define a set function ρ on $\mathcal{A} \times \mathcal{B}$ by

$$\rho(E) = \int_X \nu(E_x)\, d\mu(x).$$

Then $\rho(\varnothing) = 0$, and ρ is nonnegative. On a rectangle $A \times B$, we have

$$\rho(A \times B) = \mu(A)\nu(B) \tag{$*$}$$

since $\nu((A \times B)_x) = \nu(B) I_A$. We shall show that ρ is completely additive. If $\{E_n\}$ is a disjoint sequence in $\mathcal{A} \times \mathcal{B}$, then

$$\begin{aligned}
\rho\Big(\bigcup_n E_n\Big) &= \int_X \nu\Big(\Big(\bigcup_n E_n\Big)_x\Big)\,d\mu(x) && \text{by definition of } \rho \\
&= \int_X \nu\Big(\bigcup_n (E_n)_x\Big)\,d\mu(x) && \text{by Lemma 5.41a} \\
&= \int_X \Big[\sum_n \nu((E_n)_x)\Big]\,d\mu(x) && \text{since the sets } (E_n)_x \text{ are disjoint} \\
& && \text{for each fixed } x \\
&= \sum_n \int_X \nu((E_n)_x)\,d\mu(x) && \text{by Corollary 5.27} \\
&= \sum_n \rho(E_n).
\end{aligned}$$

Now $X \times Y = \bigcup_{m,n} (X_m \times Y_n)$. Since ρ has just been shown to be completely additive and since μ and ν are σ-finite, $(*)$ shows that ρ is σ-finite. Also, $(*)$ completely determines ρ on the algebra $\mathcal{C}$ of finite disjoint unions of rectangles. By the Extension Theorem (Theorem 5.5), ρ is completely determined on the smallest σ-algebra $\mathcal{A} \times \mathcal{B}$ containing $\mathcal{C}$.

Defining $\sigma(E) = \int_Y \mu(E^y)\,d\nu(y)$ and arguing in the same way, we see that σ is a measure on $\mathcal{A} \times \mathcal{B}$ agreeing with ρ on rectangles and determined on $\mathcal{A} \times \mathcal{B}$ by its values on rectangles. Thus we have $\rho = \sigma$ on $\mathcal{A} \times \mathcal{B}$, and can define $\mu \times \nu = \rho = \sigma$ to complete the proof.

Lemma 5.46. If f is a measurable function defined on a product space $X \times Y$, then for each x in X, $y \mapsto f(x, y)$ is a measurable function on Y, and for each y in Y, $x \mapsto f(x, y)$ is a measurable function on X.

PROOF. For each fixed x, the formula

$$\big\{y \bigm| f(x, y) > c\big\} = \big\{(x, y) \bigm| f(x, y) > c\big\}_x$$

exhibits the set on the left as a section of a measurable set, which must be measurable according to Proposition 5.42. The result for fixed y is proved similarly.

Theorem 5.47 (Fubini's Theorem). Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let $(X \times Y, \mathcal{A} \times \mathcal{B}, \mu \times \nu)$ be the product measure space. If f is a *nonnegative measurable* function on $X \times Y$, then $\int_Y f(x, y)\,d\nu(y)$ and $\int_X f(x, y)\,d\mu(x)$ are measurable, and

$$\begin{aligned}
\int_{X\times Y} f\,d(\mu \times \nu) &= \int_X \Big[\int_Y f(x, y)\,d\nu(y)\Big]\,d\mu(x) \\
&= \int_Y \Big[\int_X f(x, y)\,d\mu(x)\Big]\,d\nu(y).
\end{aligned}$$

PROOF. Lemma 5.46 shows that $f(x, y)$ is measurable in each variable separately and hence that the inside integrals in the conclusion are well defined. If f is the indicator function of a measurable subset E of $X \times Y$, then the theorem reduces to Proposition 5.45. The result immediately extends to the case of a simple function $f \geq 0$.

Now let f be an arbitrary nonnegative measurable function. Find by Proposition 5.11 an increasing sequence of simple functions $s_n \geq 0$ with pointwise limit f. The sequence of functions $\int_Y s_n(x, y)\, d\nu(y)$ is an increasing sequence of nonnegative functions, and each is measurable by what we have already shown for simple functions. By the Monotone Convergence Theorem (Theorem 5.25),

$$\lim_n \int_Y s_n(x, y)\, d\nu(y) = \int_Y \lim_n s_n(x, y)\, d\nu(y) = \int_Y f(x, y)\, d\nu(y).$$

Therefore $\int_Y f(x, y)\, d\nu(y)$ is the pointwise limit of measurable functions and is measurable. Similarly $\int_X f(x, y)\, d\mu(x)$ is measurable.

For every n, the result for simple functions gives

$$\int_{X \times Y} s_n\, d(\mu \times \nu) = \int_X \Big[\int_Y s_n(x, y)\, d\nu(y) \Big] d\mu(x).$$

By a second application of monotone convergence,

$$\int_{X \times Y} f\, d(\mu \times \nu) = \lim_n \int_{X \times Y} s_n\, d(\mu \times \nu) = \lim_n \int_X \Big[\int_Y s_n(x, y)\, d\nu(y) \Big] d\mu(x).$$

By a third application of monotone convergence,

$$\lim_n \int_X \Big[\int_Y s_n(x, y)\, d\nu(y) \Big] d\mu(x) = \int_X \Big[\lim_n \int_Y s_n(x, y)\, d\nu(y) \Big] d\mu(x).$$

Putting our results together, we obtain

$$\int_{X \times Y} f\, d(\mu \times \nu) = \int_X \Big[\int_Y f(x, y)\, d\nu(y) \Big] d\mu(x).$$

The other equality of the conclusion follows by interchanging the roles of X and Y.

Fubini's Theorem arises surprisingly often in practice. In some applications the theorem is applied at least in part to prove that an integral with a parameter is finite or is 0 for almost every value of the parameter. Here is a general result concerning integral 0.

Corollary 5.48. Suppose that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are σ-finite measure spaces, and suppose that E is a measurable subset of $X \times Y$ such that

$$\nu\big(\{y \mid (x, y) \in E\}\big) = 0$$

for almost every x $[d\mu]$. Then $\mu\big(\{x \mid (x, y) \in E\}\big) = 0$ for almost every y $[d\nu]$.

REMARKS. In words, if the x section of E has ν measure 0 for almost every x in X, then the y section of E has μ measure 0 for almost every y in Y. For example, if one-point sets in X and Y have measure 0 and if every x section of E is a finite subset of Y, then for almost every y in Y, the y section of E has measure 0 in X.

PROOF. Apply Fubini's Theorem to I_E. The iterated integrals are equal, and the hypothesis makes one of them be 0. Then the other one must be 0, and the conclusion follows.

When one tries to drop the hypothesis in Fubini's Theorem that the integrand is nonnegative, some finiteness condition is needed, and the result in the form of Theorem 5.47 is often used to establish this finiteness. Specifically suppose that f is measurable with respect to $\mathcal{A} \times \mathcal{B}$ but is not necessarily nonnegative. The assumption will be that one of the iterated integrals

$$\int_X \Big[\int_Y |f(x, y)|\, d\nu(y)\Big]\, d\mu(x) \quad \text{and} \quad \int_Y \Big[\int_X |f(x, y)|\, d\mu(x)\Big]\, d\nu(y)$$

is finite. Then the conclusions are that

(a) f is integrable with respect to $\mu \times \nu$;
(b) $\int_Y f(x, y)\, d\nu(y)$ is defined for almost every x $[d\mu]$; if it is redefined to be 0 on the exceptional set, then it is measurable and is in fact integrable $[d\mu]$;
(c) a similar conclusion is valid for $\int_X f(x, y)\, d\mu(x)$;
(d) after the redefinitions in (b) and (c), the double integral equals each iterated integral, and the two iterated integrals are equal.

These conclusions follow immediately by applying Fubini's Theorem to f^+ and f^- separately and subtracting. The redefinitions in (b) and (c) are what make the subtractions of integrands everywhere defined.

One final remark is in order: The completion of $\mathcal{A} \times \mathcal{B}$ is not necessarily the same as the product of the completions of $\mathcal{A}$ and $\mathcal{B}$, and thus the statement of Fubini's Theorem requires some modification if completions of measure spaces are to be used. We shall see in the next chapter that Borel sets in Euclidean space behave well under the formation of product spaces, but Lebesgue measurable sets do not. Thus it simplifies matters to stick to integration of Borel-measurable sets in Euclidean space whenever possible.

8. Integration of Complex-Valued and Vector-Valued Functions

Fix a measure space $(X, \mathcal{A}, \mu)$. In this chapter we have worked so far with measurable functions on X whose values are in $\mathbb{R}^*$, dividing them into two classes as far as integration is concerned. One class consists of measurable functions with values in $[0, +\infty]$, and we defined the integral of any such function as a member of $[0, +\infty]$. The other class consists of general measurable functions with values in $\mathbb{R}^*$. The integral in this case can end up being anything in $\mathbb{R}^*$, and there are some such functions for which the integral is not defined.

It is important in the theory to be able to integrate functions whose values are complex numbers or vectors in $\mathbb{R}^m$ or $\mathbb{C}^m$, and it will not be productive to allow the same broad treatment of infinities as was done for general functions with values in $\mathbb{R}^*$. On the other hand, it is desirable to have the flexibility with nonnegative measurable functions of being able to treat infinite values and infinite integrals in the same way as finite values and finite integrals. In order to have two theories, rather than three, once we pass to vector-valued functions, we shall restrict somewhat the theory we have already developed for general functions with values in $\mathbb{R}^*$.

Let us label these two theories of integration as the one for scalar-valued nonnegative measurable functions and the one for *integrable* vector-valued functions. The first of these theories has already been established and needs no change. The second of these theories needs some definitions and comments that in part repeat steps taken with Riemann integration in Sections I.5, III.3, and III.7 and in part are new. In applications of this second theory later, if the term "vector-valued" is not included in a reference to a function either explicitly or by implication, the convention is that the function is scalar-valued.

In the theory for vector-valued functions, we shall be assuming integrability, and the integrability will force the function to have meaningful finite values almost everywhere. Our convention will be that the values are finite *everywhere*. This will not be a serious restriction for any function that can be considered integrable, since we can redefine such a function on a certain set of measure 0 to be 0, and then the condition will be met without any changes in the values of integrals.

Thus let a function $f : X \to \mathbb{C}^m$ be given. Since the function can have its image contained in $\mathbb{R}^m$, we will be handling $\mathbb{R}^m$-valued functions at the same time. Since m can be 1, we will be handling complex-valued functions at the same time. Since the image can be in $\mathbb{R}^m$ and m can be 1, we will at the same time be recasting our theory of real-valued functions whose values are not necessarily nonnegative. We impose the usual Hermitian inner product $(\cdot\,,\,\cdot)$ and norm $|\cdot|$ on $\mathbb{C}^m$.

The function $\bar{f} : X \to \mathbb{C}^m$ is the composition of f followed by complex conjugation in each entry of $\mathbb{C}^m$. We can write $f = \operatorname{Re} f + i \operatorname{Im} f$, where

$\operatorname{Re} f = \frac{1}{2}(f + \bar{f})$ and $\operatorname{Im} f = \frac{1}{2i}(f - \bar{f})$, and then the functions $\operatorname{Re} f$ and $\operatorname{Im} f$ take values in $\mathbb{R}^m$. Following the convention in Section A7 of the appendix, let $\{u_1, \dots, u_m\}$ be the standard basis of $\mathbb{R}^m$.

By a basic open set in $\mathbb{C}^m$, we mean a set that is a product in $\mathbb{R}^{2m}$ of bounded open intervals in each coordinate. In symbols, such a set is centered at some $v_0 \in \mathbb{C}^m$, and there are positive numbers ξ_j and η_j such that the set is

$$\left\{v \in \mathbb{C}^m \;\middle|\; |(\operatorname{Re}(v - v_0), u_j)| < \xi_j \text{ and } |(\operatorname{Im}(v - v_0), u_j)| < \eta_j \text{ for } 1 \le j \le m\right\}.$$

We say that $f : X \to \mathbb{C}^m$ is **measurable** if the inverse image under f of each basic open set in $\mathbb{C}^m$ is measurable, i.e., lies in $\mathcal{A}$.

Lemma 5.49. A function $f : X \to \mathbb{C}^m$ is measurable if and only if the inverse image under f of each open set in $\mathbb{C}^m$ is in $\mathcal{A}$.

PROOF. If the stated condition holds, then the inverse image of any *basic* open set is in $\mathcal{A}$, and hence f is measurable. Conversely suppose f is measurable, and let an open set U in $\mathbb{C}^m$ be given. Then U is the union of a sequence of basic open sets U_n, and the measurability of f, in combination with the formula $f^{-1}(U) = \bigcup_n f^{-1}(U_n)$, shows that $f^{-1}(U)$ is in $\mathcal{A}$.

Proposition 5.50. A function $f : X \to \mathbb{C}^m$ is measurable if and only if $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable.

PROOF. In view of Lemma 5.49, we can work with arbitrary open sets in place of basic open sets. If U and V are open sets in $\mathbb{R}^m$, then the product set $U + iV$ is open in $\mathbb{C}^m$, and $f^{-1}(U + iV) = (\operatorname{Re} f)^{-1}(U) \cap (\operatorname{Im} f)^{-1}(V)$. It is immediate that measurability of $\operatorname{Re} f$ and $\operatorname{Im} f$ implies measurability of f. Conversely if we specialize this formula to $V = \mathbb{R}^m$, then we see that measurability of f implies measurability of $\operatorname{Re} f$. Similarly if we specialize to $U = \mathbb{R}^m$, then we see that measurability of f implies measurability of $\operatorname{Im} f$.

Proposition 5.51. The following conditions on a function $f : X \to \mathbb{C}^m$ are equivalent:

(a) f is measurable,
(b) (f, v) is measurable for each v in $\mathbb{C}^m$,
(c) (f, u_j) is measurable for $1 \le j \le m$.

REMARKS. When infinite-dimensional ranges are used in more advanced texts, (a) is summarized by saying that f is "strongly measurable," and (b) is summarized by saying that f is "weakly measurable."

PROOF. Suppose (a) holds. The function in (b) is the composition of f followed by the continuous function $(\,\cdot\,, v)$ from $\mathbb{C}^m$ to $\mathbb{C}$. The inverse image of an open set in $\mathbb{C}$ is then open in $\mathbb{C}^m$, and the inverse image of the latter open set under f is in $\mathcal{A}$. This proves (b). Condition (b) trivially implies condition (c). If (c) holds, then Proposition 5.50 shows that $(\operatorname{Re} f, u_j)$ and $(\operatorname{Im} f, u_j)$ are measurable from X into $\mathbb{R}$. Thus the inverse image of any open interval under any of these $2m$ functions on X is in $\mathcal{A}$. The inverse image of a basic open set in $\mathbb{C}^m$ under f is the intersection of $2m$ such sets in $\mathcal{A}$ and is therefore in $\mathcal{A}$. Hence (a) holds.

Proposition 5.52. Measurability of vector-valued functions has the following properties:

(a) If $f : X \to \mathbb{C}^m$ and $g : X \to \mathbb{C}^m$ are measurable, then so is $f + g$ as a function from X to $\mathbb{C}^m$.

(b) If $f : X \to \mathbb{C}^m$ is measurable and c is in $\mathbb{C}$, then cf is measurable as a function from X to $\mathbb{C}^m$.

(c) If $f : X \to \mathbb{C}^m$ is measurable, then so is $\bar{f} : X \to \mathbb{C}^m$.

(d) If $f : X \to \mathbb{C}$ and $g : X \to \mathbb{C}$ are measurable, then so is $fg : X \to \mathbb{C}$.

(e) If $f : X \to \mathbb{C}^m$ is measurable, then $|f| : X \to [0, +\infty)$ is measurable.

(f) If $\{f_n\}$ is a sequence of measurable functions from X into $\mathbb{C}^m$ converging pointwise to a function $f : X \to \mathbb{C}^m$, then f is measurable.

PROOF. Conclusions (a) through (e) may all be proved in the same way. It will be enough to illustrate the technique with (a). We can write the function $x \mapsto f(x) + g(x)$ as a composition of $x \mapsto (f(x), g(x))$ followed by addition $(a, b) \mapsto a+b$. Let an open set in $\mathbb{C}^m$ be given. The inverse image under addition is open in $\mathbb{C}^m \times \mathbb{C}^m$, since addition is continuous (Proposition 2.28). The inverse image of a product $U \times V$ of open sets in $\mathbb{C}^m \times \mathbb{C}^m$ under $x \mapsto (f(x), g(x))$ is $f^{-1}(U) \cap g^{-1}(V)$, which is in $\mathcal{A}$ because f and g are measurable, and therefore the inverse image of any open set in $\mathbb{C}^m \times \mathbb{C}^m$ under $x \mapsto (f(x), g(x))$ is in $\mathcal{A}$. This handles (a), and (b) through (e) are similar.

For (f), we apply Proposition 5.50 to f, and then we apply the equivalence of (a) and (c) of Proposition 5.51 for $\operatorname{Re} f$ and $\operatorname{Im} f$. In this way the result is reduced to the real-valued scalar case, which is known from Corollary 5.10.

If E is a measurable subset of X, we say that a function $f : X \to \mathbb{C}$ is **integrable** on E if $\operatorname{Re} f$ and $\operatorname{Im} f$ are integrable on E, and in this case we define $\int_E f\,d\mu = \int_E \operatorname{Re} f\,d\mu + i \int_E \operatorname{Im} f\,d\mu$.

Proposition 5.53. Let E be a measurable subset of X. Integrability on E of functions from X to $\mathbb{C}$ has the following properties:

(a) If f and g are functions from X into $\mathbb{C}$ that are integrable on E, then $f+g$ is integrable on E, and $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.

(b) If f is a function from X into $\mathbb{C}$ that is integrable on E and if c is in $\mathbb{C}$, then cf is integrable on E, and $\int_E cf\,d\mu = c\int_E f\,d\mu$.

(c) If f is a measurable function from X into $\mathbb{C}$ such that $|f|$ is integrable on E, then f is integrable on E, and $\big|\int_E f(x)\,d\mu(x)\big| \le \int_E |f(x)|\,d\mu(x)$.

(d) (Dominated convergence) Let f_n be a sequence of measurable functions from X into $\mathbb{C}$ integrable on E and converging pointwise to f. If there is a measurable function $g : X \to [0, +\infty]$ that is integrable on E and has $|f_n(x)| \le g(x)$ for all x in E, then f is integrable on E, $\lim_n \int_E f_n\,d\mu$ exists in $\mathbb{C}$, and $\lim_n \int_E f_n\,d\mu = \int_E f\,d\mu$.

PROOF. Conclusion (a) is immediate from the definitions, and so is (b) for real scalars. Taking (a) and (b) into account, we see that (b) holds if it holds for $c = i$. We have $if = -\operatorname{Im} f + i\operatorname{Re} f$. If f is integrable, then $-\operatorname{Im} f$ and $\operatorname{Re} f$ are integrable, and hence if is integrable. Then

$$\begin{aligned} i\int_E f\,d\mu &= i\big(\int_E \operatorname{Re} f\,d\mu + i\int_E \operatorname{Im} f\,d\mu\big) \\ &= \int_E (-\operatorname{Im} f)\,d\mu + \int_E (i\operatorname{Re} f)\,d\mu = \int_E if\,d\mu, \end{aligned}$$

and hence (b) is proved.

In (c), if $f : X \to \mathbb{C}$ is integrable, choose c with $|c| = 1$ such that $c\int_E f\,d\mu$ is real and ≥ 0. Application of (b) and Proposition 5.16 gives $\big|\int_E f\,d\mu\big| = c\int_E f\,d\mu = \int_E cf\,d\mu = \int_E \operatorname{Re}(cf)\,d\mu \le \int_E |cf|\,d\mu = \int_E |f|\,d\mu$.

Finally (d) follows by applying the Dominated Convergence Theorem (Theorem 5.30) to $\operatorname{Re} f_n$ and $\operatorname{Im} f_n$ separately and then combining the results.

We turn now to the matter of integrability of vector-valued functions, together with the value of the integral. One way of proceeding is to go back and adapt the theory in Sections 3–4 to work directly with vector-valued functions and approximations by vector-valued simple functions. This approach is useful if at some stage one wants systematically to allow infinite-dimensional vectors as values. Examples of this situation will arise in this book, but there are not enough examples to justify an abstract treatment. One important example arises in the next section with functions of the form $f(x, y)$, which can be regarded as functions of x that take values in a space of functions of y.

Thus we use an abstract definition of integrability that is appropriate only to the case of finite-dimensional range. If E is a measurable subset of X, we say that a function $f : X \to \mathbb{C}^m$ is **integrable** on E if the complex-valued functions (f, u_j) are integrable on E for each u_j in the standard basis, and in this case we define $\int_E f\,d\mu = \sum_{j=1}^n \big(\int_E (f, u_j)\,d\mu\big)u_j$.

Proposition 5.54. Let E be a measurable subset of X. Integrability of vector-valued functions on E satisfies the following properties:

(a) If f and g are functions from X into $\mathbb{C}^m$ that are integrable on E, then $f+g$ is integrable on E, and $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.

(b) If f is a function from X into $\mathbb{C}^m$ that is integrable on E, then cf is integrable on E, and $\int_E cf\,d\mu = c\int_E f\,d\mu$.

(c) A function $f : X \to \mathbb{C}^m$ is integrable on E if and only if $\operatorname{Re} f$ and $\operatorname{Im} f$ are integrable on E, and then $\int_X f\,d\mu = \int_X \operatorname{Re} f\,d\mu + i\int_X \operatorname{Im} f\,d\mu$.

(d) If f is a function from X into $\mathbb{C}^m$ that is integrable on E and if v is a member of $\mathbb{C}^m$, then $x \mapsto (f(x), v)$ is integrable on E and $\int_E (f(x), v)\,d\mu(x) = \big(\int_E f(x)\,d\mu(x), v\big)$.

(e) If f is a measurable function from X into $\mathbb{C}^m$ such that $|f|$ is integrable on E, then f is integrable on E, and $\big|\int_E f(x)\,d\mu(x)\big| \le \int_E |f(x)|\,d\mu(x)$.

(f) (Dominated convergence) Let f_n be a sequence of measurable functions from X into $\mathbb{C}^m$ integrable on E and converging pointwise to f. If there is a measurable function $g : X \to [0, +\infty]$ that is integrable on E and has $|f_n(x)| \le g(x)$ for all x in E, then f is integrable on E, $\lim_n \int_E f_n\,d\mu$ exists in $\mathbb{C}^m$, and $\lim_n \int_E f_n\,d\mu = \int_E f\,d\mu$.

PROOF. All of the relevant questions about measurability are addressed by Propositions 5.50 and 5.52. Conclusions (a), (b), (c), and (f) about integrability are immediate from Proposition 5.53.

For (d), let $v = \sum c_j u_j$ with each c_j in $\mathbb{C}$. Since f is by assumption integrable, $(f, v) = (f, \sum c_j u_j) = \sum_j \bar{c}_j (f, u_j)$ exhibits (f, v) as a linear combination of functions integrable on E. Therefore (f, v) is integrable on E. To obtain the formula asserted in (d), we first consider $v = u_i$. Then the definition of $\int_E f\,d\mu$ gives $\big(\int_E f\,d\mu, u_i\big) = \big(\sum_j \big(\int_E (f, u_j)\,d\mu\big) u_j,\, u_i\big) = \int_E (f, u_i)\,d\mu$. Multiplying by $\bar{c}_i$ and adding, we obtain $\big(\int_E f\,d\mu, v\big) = \int_E (f, v)\,d\mu$. This proves (d).

For (e), let $f : X \to \mathbb{C}^m$ be integrable on E. The asserted inequality is trivial if $\int_E f\,d\mu = 0$. Otherwise, for every v in $\mathbb{C}^m$,

$$\begin{aligned}
\big|\big(\textstyle\int_E f\,d\mu, v\big)\big| &= \big|\textstyle\int_E (f, v)\,d\mu\big| && \text{by (d)}\\
&\le \textstyle\int_E |(f, v)|\,d\mu && \text{by Proposition 5.53c}\\
&\le |v| \textstyle\int_E |f|\,d\mu && \text{by Proposition 5.16 and the Schwarz inequality.}
\end{aligned}$$

Taking $v = \int_E f\,d\mu$ gives $\big|\int_E f\,d\mu\big|^2 \le \big|\int_E f\,d\mu\big| \int_E |f|\,d\mu$. Since $\int_E f\,d\mu$ has been assumed nonzero, we can divide by its magnitude, and then (e) follows.

9. L^1, L^2, L^∞, and Normed Linear Spaces

Let $(X, \mathcal{A}, \mu)$ be a measure space. In this section we introduce the spaces $L^1(X)$, $L^2(X)$, and $L^\infty(X)$. Roughly speaking, these will be vector spaces of functions on X with suitable integrability properties. More precisely the actual vector spaces of functions will form pseudometric spaces, and the spaces $L^1(X)$, $L^2(X)$, and $L^\infty(X)$ will be the corresponding metric spaces obtained from the construction of Proposition 2.12. They will all turn out to be vector spaces over $\mathbb{R}$ or $\mathbb{C}$. It will matter little whether the scalars for these vector spaces are real or complex. When we need to refer to operations with scalars, we may use the symbol $\mathbb{F}$ to denote $\mathbb{R}$ or $\mathbb{C}$, and we call $\mathbb{F}$ the **field of scalars**. We shall make explicit mention of $\mathbb{R}$ or $\mathbb{C}$ in any situation in which it is necessary to insist on a particular one of $\mathbb{R}$ or $\mathbb{C}$.

The three spaces we will construct will all be obtained by introducing "pseudonorms" in vector spaces of measurable functions. A **pseudonorm** on a vector space V is a function $\|\cdot\|$ from V to $[0, +\infty)$ such that[7]

(i) $\|x\| \geq 0$ for all $x \in V$,

(ii) $\|cx\| = |c|\|x\|$ for all scalars c and all $x \in V$,

(iii) **(triangle inequality)** $\|x + y\| \leq \|x\| + \|y\|$ for all x and y in V.

We encountered pseudonorms earlier in connection with pseudo inner-product spaces; in Proposition 2.3 we saw how to form a pseudonorm from a pseudo inner product. However, only the pseudonorm for $L^2(X)$ arises from a pseudo inner product in the construction of L^1, L^2, and L^∞.

The definitions of the pseudonorms in these three instances are

$$\begin{aligned} \|f\|_1 &= \textstyle\int_X |f|\,d\mu && \text{for } L^1(X), \\ \|f\|_2 &= \left(\textstyle\int_X |f|^2\,d\mu\right)^{1/2} && \text{for } L^2(X), \\ \|f\|_\infty &= \text{“essential supremum” of } f && \text{for } L^\infty(X). \end{aligned}$$

Once we have defined "essential supremum," all the above expressions are meaningful for any measurable function f from X to the scalars, and the vector space V in each of the cases is the space of all measurable functions from X to the scalars such that the indicated pseudonorm is finite. In other words, V consists of the integrable functions on X in the case of $L^1(X)$, the square-integrable functions on X in the case of $L^2(X)$, and the "essentially bounded" functions on X in the case of $L^\infty(X)$.

We need to check that $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ are indeed pseudonorms and that the spaces V are vector spaces in each case.

---

[7] The word "seminorm" is a second name for a function with these properties and is generally used in the context of a family of such functions. We shall not use the word "seminorm" in this text.

For $L^1(X)$, properties (i) and (ii) are immediate from the definition. For (iii), we have $|f(x)+g(x)| \le |f(x)|+|g(x)|$ for all x and therefore $\|f+g\|_1 = \int_X |f+g|\,d\mu \le \int_X |f|\,d\mu + \int_X |g|\,d\mu = \|f\|_1 + \|g\|_1$.

For $L^2(X)$, let V be the space of all square-integrable functions on X. The space V is certainly closed under scalar multiplication; let us see that it is closed under addition. If f and g are in V, then we have

$$\begin{aligned}(|f(x)|+|g(x)|)^2 &\le \big(\max\{|f(x)|,|g(x)|\}+\max\{|f(x)|,|g(x)|\}\big)^2\\ &= 4\max\{|f(x)|^2,|g(x)|^2\} \le 4|f(x)|^2+4|g(x)|^2\end{aligned}$$

for every x in X. Integrating over X, we see that $f+g$ is in V if f and g are in V. Also, the left side is $\ge 4|f(x)|\,|g(x)|$, and it follows that $f\bar g$ is integrable whenever f and g are in V. Then the definition $(f,g)_2 = \int_E f\bar g\,d\mu$ makes V into a pseudo inner product-space in the sense of Section II.1. Hence Proposition 2.3 shows that the function $\|\cdot\|_2$ with $\|f\|_2 = (f,f)_2^{1/2}$ is a pseudonorm on V.

For $L^\infty(X)$, we say that f is **essentially bounded** if there is a real number M such that $|f(x)| \le M$ almost everywhere $[d\mu]$. Let us call such an M an **essential bound** for $|f|$. When f is essentially bounded, we define $\|f\|_\infty$ to be the infimum of all essential bounds for $|f|$. This infimum is itself an essential bound, since the countable union of sets of measure 0 is of measure 0. The infimum of the essential bounds is called the **essential supremum** of $|f|$. Certainly $\|\cdot\|_\infty$ satisfies (i) and (ii). If $|f|$ is bounded a.e. by M and if $|g|$ is bounded a.e. by N, then $|f+g|$ is bounded everywhere by $|f|+|g|$, which is bounded a.e. by $M+N$. It follows that $f+g$ is essentially bounded and $\|f+g\|_\infty \le \||f|+|g|\|_\infty \le \|f\|_\infty + \|g\|_\infty$. So (iii) holds for $\|\cdot\|_\infty$.

A real or complex vector space with a pseudonorm is a **pseudo normed linear space**. Such a space V becomes a pseudometric space by the definition $d(f,g) = \|f-g\|$, according to the proof of Proposition 2.3. Proposition 2.12 shows that if we define two members f and g of V to be equivalent whenever $d(f,g)=0$, then the result is an equivalence relation and the function d descends to a well-defined metric on the set of equivalence classes. If we take into account the vector space structure on V, then we can see that the operations of addition and scalar multiplication descend to the set of equivalence classes, and the set of equivalence classes is then also a vector space. The argument for addition is that if $d(f_1,f_2)=0$ and $d(g_1,g_2)=0$, then $d(f_1+g_1,f_2+g_2)$ is 0 because

$$\begin{aligned}d(f_1+g_1,f_2+g_2) &= \|(f_1+g_1)-(f_2+g_2)\| = \|(f_1-f_2)+(g_1-g_2)\|\\ &\le \|f_1-f_2\|+\|g_1-g_2\| = d(f_1,f_2)+d(g_1,g_2) = 0.\end{aligned}$$

The argument for scalar multiplication is similar, and one readily checks that the space of equivalence classes is a vector space.

This construction is to be applied to the spaces V we formed in connection with integrability, square integrability, and essential boundedness. The spaces of equivalence classes in the respective cases are called $L^1(X)$, $L^2(X)$, and $L^\infty(X)$. These spaces of equivalence classes are pseudo normed linear spaces with the additional property that $\|f\| = 0$ only for the 0 element of the vector space. If there is any possibility of confusion, we may write $L^1(\mu)$ or $L^1(X, \mu)$ or $L^1(X, \mathcal{A}, \mu)$ in place of $L^1(X)$, and similarly for L^2 and L^∞.

A pseudo normed linear space is called a **normed linear space** if $\|f\| = 0$ implies f is the 0 element of the vector space. Thus $L^1(X)$, $L^2(X)$, and $L^\infty(X)$ are normed linear spaces.

In practice, in order to avoid clumsiness, one sometimes relaxes the terminology and works with the members of $L^1(X)$, $L^2(X)$, and $L^\infty(X)$ as if they were functions, saying, "Let the function f be in $L^1(X)$" or "Let f be an L^1 function." There is little possibility of ambiguity in using such expressions.

The 1-dimensional vector space consisting of the field of scalars $\mathbb{F}$ with absolute value as norm is an example of a normed linear space. Apart from this and $\mathbb{F}^m$, we have encountered one other important normed linear space thus far in the book. This is the space $B(S)$ of bounded functions on a nonempty set S. It has various vector subspaces of interest, such as the space $C(S)$ of bounded continuous functions in the case that S is a metric space. The norm for $B(S)$ is the **supremum norm** or the **uniform norm** defined by

$$\|f\|_{\sup} = \sup_{s\in S} |f(s)|.$$

The corresponding metric is

$$d(f, g) = \|f - g\|_{\sup} = \sup_{s\in S} |f(s) - g(s)|,$$

and this agrees with the definition of the metric in the example in Chapter II. Proposition 2.44 shows that the metric space $B(S)$ is complete. Any vector subspace of $B(S)$ is a normed linear space under the restriction of the supremum norm to the subspace.

In working with specific normed linear spaces, we shall often be interested in seeing whether a particular subset of the space is dense. In checking denseness, the following proposition about an arbitrary normed linear space is sometimes helpful. The intersection of vector subspaces of X is a vector subspace, and the intersection of closed sets is closed. Therefore it makes sense to speak of the **smallest closed vector subspace** containing a given subset S of X.

Proposition 5.55. If X is a normed linear space with norm $\|\cdot\|$ and with $\mathbb{F}$ as field of scalars, then

(a) addition is a continuous function from $X \times X$ to X,
(b) scalar multiplication is a continuous function from $\mathbb{F} \times X$ to X,

(c) the closure of any vector subspace of X is a vector subspace,
(d) the set of all finite linear combinations of members of a subset S of X is dense in the smallest closed vector subspace containing S.

PROOF. The formula $\|(x+y)-(x_0+y_0)\| \le \|x-x_0\| + \|y-y_0\|$ shows continuity of addition because it says that if x is within distance $\epsilon/2$ of x_0 and y is within distance $\epsilon/2$ of y_0, then $x+y$ is within distance ϵ of x_0+y_0. Similarly the formula $\|cx-c_0x_0\| \le \|c(x-x_0)\| + \|(c-c_0)x_0\| = |c|\|x-x_0\| + |c-c_0|\|x_0\|$ shows that $\|cx-c_0x_0\| \le \delta(|c_0|+1)+\delta\|x_0\|$ as soon as $\delta \le 1$, $|c-c_0| \le \delta$, and $\|x-x_0\| \le \delta$. If ϵ with $0<\epsilon\le 1$ is given and if we set $\delta = (|c_0|+1+\|x_0\|)^{-1}\epsilon$, then we see that $|c-c_0| \le \delta$ and $\|x-x_0\| \le \delta$ together imply $\|cx-c_0x_0\| \le \epsilon$. Hence scalar multiplication is continuous. This proves (a) and (b).

From (a) and (b) it follows that if $x_n \to x$ and $y_n \to y$ in X and $c_n \to c$ in $\mathbb{F}$, then $x_n + y_n \to x+y$ and $c_nx_n \to cx$. This proves (c).

For (d), the smallest closed vector subspace V_1 containing S certainly contains the closure V_2 of the set of all finite linear combinations of members of S. Part (c) shows that V_2 is a closed vector subspace, and hence the definition of V_1 implies that V_1 is contained in V_2. Therefore $V_1 = V_2$, and (d) is proved.

Proposition 5.56. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $p=1$ or $p=2$. Then every indicator function of a set of finite measure is in $L^p(X)$, and the smallest closed subspace of $L^p(X)$ containing all such indicator functions is $L^p(X)$ itself.

REMARK. Proposition 5.55d allows us to conclude from this that the the set of simple functions built from sets of finite measure lies in both $L^1(X)$ and $L^2(X)$ and is dense in each. It of course lies in $L^\infty(X)$ as well, but it is dense in $L^\infty(X)$ if and only if $\mu(X)$ is finite.

PROOF. If E is a set of finite measure, then the equality $\int_X (I_E)^p\,d\mu = \mu(E)$ shows that I_E is in L^p for $p=1$ and $p=2$.

In the reverse direction let V be the smallest closed vector subspace of L^p containing all indicator functions of sets of finite measure. Suppose that $s = \sum_k c_k I_{E_k}$ is the canonical expansion of a simple function $s \ge 0$ in L^p and that $c_k > 0$. The inequalities $0 \le c_k I_{E_k} \le s$ imply that $c_k I_{E_k}$ is in L^p. Hence I_{E_k} is in L^p, and $\mu(E_k)$ is finite. Thus every nonnegative simple function in L^p lies in V.

Let $f \ge 0$ be in L^p, and let s_n be an increasing sequence of simple functions ≥ 0 with pointwise limit f. Since $0 \le s_n \le f$, each s_n is in L^p. Since $|f-s_n|^p$ has pointwise limit 0 and is dominated pointwise for every n by the integrable function $|f|^p$, dominated convergence gives $\lim \int_X |f-s_n|^p\,d\mu = 0$. Hence s_n tends to f in L^p. Combining this conclusion with the result of the previous paragraph, we see that every nonnegative L^p function is in V. Any L^p function is a finite linear combination of nonnegative L^p functions, and hence every L^p function lies in V.

Let us digress briefly once more from our study of L^1, L^2, and L^∞ to obtain two more results about general normed linear spaces. A linear function between two normed linear spaces is often called a **linear operator**. A linear function whose range space is the field of scalars is called a **linear functional**. The following equivalence of properties is fundamental and is often used without specific reference.

Proposition 5.57. Let X and Y be normed linear spaces that are both real or both complex, and let their respective norms be $\|\cdot\|_X$ and $\|\cdot\|_Y$. Then the following conditions on a linear operator $L : X \to Y$ are equivalent:

(a) L is uniformly continuous on X,
(b) L is continuous on X,
(c) L is continuous at 0,
(d) L is **bounded** in the sense that there exists a constant M such that

$$\|L(x)\|_Y \le M\|x\|_X$$

for all x in X.

PROOF. If L is uniformly continuous on X, then L is certainly continuous on X. If L is continuous on X, then L is certainly continuous at 0. Thus (a) implies (b), and (b) implies (c).

If L is continuous at 0, find $\delta > 0$ for $\epsilon = 1$ such that $\|x - 0\|_X \le \delta$ implies $\|L(x) - L(0)\|_Y \le 1$. Here $L(0) = 0$. If a general $x \ne 0$ is given, then $\|x\|_X \ne 0$, and the properties of the norm give $\|(\delta/\|x\|_X)x\|_X = \delta$. Thus $\|L((\delta/\|x\|_X)x)\|_Y \le 1$. By the linearity of L and the properties of the norm, $(\delta/\|x\|_X)\|L(x)\|_Y \le 1$. Therefore $\|L(x)\|_Y \le \delta^{-1}\|x\|_X$, and L is bounded with $M = \delta^{-1}$. Thus (c) implies (d).

If L is bounded with constant M and if $\epsilon > 0$ is given, let $\delta = \epsilon/M$. Then $\|x_1 - x_2\|_X \le \delta$ implies

$$\|L(x_1) - L(x_2)\|_Y = \|L(x_1 - x_2)\|_Y \le M\|x_1 - x_2\|_X \le \delta M = \epsilon.$$

Thus (d) implies (a).

If $L : X \to Y$ is a bounded linear operator, then the infimum of all constants M such that $\|L(x)\|_Y \le M\|x\|_X$ for all x in X is again such a constant, and it is called the **operator norm** $\|L\|$ of L. Thus it in particular satisfies

$$\|L(x)\|_Y \le \|L\|\|x\|_X \qquad \text{for all } x \text{ in } X.$$

As a consequence of the way that L and the norms in X and Y interact with scalar multiplication, the operator norm is given by the formulas

$$\|L\| = \sup_{\|x\|_X \le 1} \|L(x)\|_Y = \sup_{\|x\|_X = 1} \|L(x)\|_Y$$

except in the uninteresting case $X = 0$. It is easy to check that the bounded linear operators from X into Y form a vector space, and the operator norm makes this vector space into a normed linear space that we denote by $\mathcal{B}(X, Y)$. When the domain and range are the same space X, we refer to the members of $\mathcal{B}(X, X)$ as bounded linear operators *on* X. The normed linear space $\mathcal{B}(X, X)$ has a multiplication operation given by composition.

When Y is the field of scalars $\mathbb{F}$, the space $\mathcal{B}(X, \mathbb{F})$ reduces to the space of continuous linear functionals on X. This is called the **dual space** of X and is denoted by X^*. For example, if $X = L^1(\mu)$, then every member g of $L^\infty(\mu)$ defines a member x_g^* of X^* by $x_g^*(f) = \int fg\,d\mu$ for f in $L^1(\mu)$; the linear functional x_g^* has $\|x_g^*\| \le \|g\|_\infty$. We shall be interested in two kinds of convergence in X^*. One is **norm convergence**, in which a sequence $\{x_n^*\}$ converges to an element x^* in X^* if $\|x_n^* - x^*\|$ tends to 0. The other is **weak-star convergence**, in which $\{x_n^*\}$ converges to x^* weak-star against X if $\lim_n x_n^*(x) = x^*(x)$ for each x in X.

Theorem 5.58 (Alaoglu's Theorem, preliminary form). If X is a separable normed linear space, then any sequence in X^* that is *bounded in norm* has a subsequence that converges weak-star against X.

REMARKS. In Chapter VI we shall see that L^1 and L^2 are separable in the case of Lebesgue measure on $\mathbb{R}^1$ and in the case of many generalizations of Lebesgue measure to N-dimensional Euclidean space.

PROOF. Let a sequence $\{x_n^*\}_{n=1}^\infty$ be given with $\|x_n^*\| \le M$, and let $\{x_k\}$ be a countable dense set in X. For each k, we have $|x_n^*(x_k)| \le \|x_n^*\|\|x_k\| \le M\|x_k\|$, and hence the sequence $\{x_n^*(x_k)\}_{n=1}^\infty$ of scalars is bounded for each fixed k. By the Bolzano–Weierstrass Theorem, $\{x_n^*(x_k)\}_{n=1}^\infty$ has a convergent subsequence. Since we can pass to a convergent subsequence of any subsequence for any particular k, we can use a diagonal process to pass to a single convergent subsequence $\{x_{n_l}^*\}_{l=1}^\infty$ such that $\lim_l x_{n_l}^*(x_k)$ exists for all k.

Now let x_0 be arbitrary in X, let $\epsilon > 0$ be given, and choose x_k in the dense set with $\|x_k - x_0\| < \epsilon$. Then

$$\begin{aligned} |x_{n_l}^*(x_0) - x_{n_{l'}}^*(x_0)| &\le |x_{n_l}^*(x_0 - x_k)| + |x_{n_l}^*(x_k) - x_{n_{l'}}^*(x_k)| + |x_{n_{l'}}^*(x_k - x_0)| \\ &\le M\|x_0 - x_k\| + |x_{n_l}^*(x_k) - x_{n_{l'}}^*(x_k)| + M\|x_k - x_0\| \\ &\le 2M\epsilon + |x_{n_l}^*(x_k) - x_{n_{l'}}^*(x_k)|. \end{aligned}$$

Thus $\limsup_{l,l'\to\infty} |x^*_{n_l}(x_0) - x^*_{n_{l'}}(x_0)| \le 2M\epsilon$. Since ϵ is arbitrary, we conclude that $\{x^*_{n_l}(x_0)\}_{l=1}^\infty$ is a Cauchy sequence of scalars. It is therefore convergent. Denote the limit by $x^*(x_0)$, so that $\lim_l x^*_{n_l}(x_0) = x^*(x_0)$ for all x_0 in X. Since limits respect addition and multiplication of scalars, x^* is a linear functional on X. The computation $|x^*(x_0)| = |\lim_l x^*_{n_l}(x_0)| = \lim_l |x^*_{n_l}(x_0)| \le \limsup_l \|x^*_l\|\|x_0\| \le M\|x_0\|$ shows that x^* is bounded. Hence $\{x^*_{n_l}\}_{l=1}^\infty$ converges to x^* weak-star against X.

Now, as promised, we return to L^1, L^2, and L^∞. The completeness asserted in the next theorem will turn out to be one of the key advantages of Lebesgue integration over Riemann integration.

Theorem 5.59. Let $(X, \mathcal{A}, \mu)$ be any measure space, and let p be 1, 2, or ∞. Any Cauchy sequence $\{f_k\}$ in L^p has a subsequence $\{f_{k_n}\}$ such that $\|f_{k_n} - f_{k_m}\|_p \le C_{\min\{m,n\}}$ with $\sum_n C_n < +\infty$. A subsequence $\{f_{k_n}\}$ with this property is necessarily Cauchy pointwise almost everywhere. If f denotes the almost-everywhere limit of $\{f_{n_k}\}$, then the original sequence $\{f_k\}$ converges to f in L^p. Consequently these three spaces L^p, when regarded as metric spaces, are complete in the sense that every Cauchy sequence converges.

REMARKS. The broad sweep of the theorem is that the spaces L^1, L^2, and L^∞ are complete. But the detail is important, too. First of all, the detail allows us to conclude that a sequence convergent in one of these spaces has an almost-everywhere convergent subsequence. Second of all, the detail allows us to conclude that if a sequence of functions is convergent in L^{p_1} and in L^{p_2}, then the limit functions in the two spaces are equal almost everywhere.

PROOF. Let $\{f_n\}$ be a Cauchy sequence in L^p. Inductively choose integers n_k by defining $n_0 = 1$ and taking n_k to be any integer $> n_{k-1}$ such that $\|f_m - f_{n_k}\|_p \le 2^{-k}$ for $m \ge n_k$; we can do so since the given sequence is Cauchy. Then the subsequence $\{f_{k_n}\}$ has the property that $\|f_m - f_n\|_p \le 2^{-\min\{m,n\}}$ for all $m \ge 1$ and $n \ge 1$. This proves the first conclusion of the theorem.

Now suppose that we have a sequence $\{f_n\}$ in L^p such that $\|f_n - f_m\|_p \le C_{\min\{m,n\}}$ with $\sum_n C_n = C < +\infty$. We shall prove that $\{f_n\}$ is Cauchy pointwise almost everywhere and that if f is its almost-everywhere limit, then f_n tends to f in L^p.

First suppose that $p < \infty$. Let g_n be the function from X to $[0, +\infty]$ given by

$$g_n = |f_1| + \sum_{k=2}^{n} |f_k - f_{k-1}|, \qquad (*)$$

and define $g(x) = \lim g_n(x)$ pointwise. Then

$$\Big(\int_X g_n\,d\mu\Big)^{1/p} = \|g_n\|_p \le \|f_1\|_p + \sum_{k=2}^{n} \|f_k - f_{k-1}\|_p$$
$$\le \|f_1\|_p + \sum_{k=2}^{n} C_{k-1} \le \|f_1\|_p + C.$$

By monotone convergence, we deduce that $\big(\int_X g\,d\mu\big)^{1/p} = \|g\|_p$ is finite. Thus g is finite a.e., and consequently the series

$$\sum_{k=2}^{\infty} |f_k(x) - f_{k-1}(x)| \qquad \text{converges in } \mathbb{R} \text{ for a.e. } x\ [d\mu]. \tag{$**$}$$

By redefining the functions f_k on a set of μ measure 0, we may assume that the series $(**)$ converges pointwise to a limit in $\mathbb{R}$ for every x. Consequently the series

$$\sum_{k=2}^{\infty} (f_k(x) - f_{k-1}(x))$$

is absolutely convergent for all x and must be convergent for all x. The partial sums for the series without the absolute value signs are $f_n(x) - f_1(x)$, and hence $f(x) = \lim f_n(x)$ exists in $\mathbb{R}$ for every x. For every n,

$$|f - f_n| \le \sum_{k=n+1}^{\infty} |f_k - f_{k-1}| \le g, \tag{$\dagger$}$$

and we have seen that g^p is integrable. By dominated convergence, we conclude that $\lim_n \int_X |f - f_n|^p\,d\mu = \int_X \lim_n |f(x) - f_n(x)|^p\,d\mu(x) = 0$. In other words, $\lim_n \|f - f_n\|_p = 0$. Therefore f_n tends to f in $L^p(\mu)$.

Next suppose that $p = \infty$. Let $\{f_n\}$ be any Cauchy sequence in L^∞. For each m and n, let E_{mn} be the subset of X where $|f_m - f_n| > \|f_m - f_n\|_\infty$, and put $E = \bigcup_{m,n} E_{mn}$. This set has measure 0. Redefine all functions to be 0 on E. The redefined functions are then uniformly Cauchy, hence uniformly convergent to some function f, and then f_n tends to f in $L^\infty(X)$.

For any p, we have shown that the original Cauchy sequence $\{f_n\}$ has a convergent subsequence $\{f_{n_k}\}$ in L^p. Let f be the L^p limit of the subsequence. Given $\epsilon > 0$, choose N such that $n \ge m \ge N$ implies $\|f_n - f_m\|_p \le \epsilon$, and then choose K such that $\|f_{n_k} - f\|_p \le \epsilon$ for $k \ge K$. Fix $k \ge K$ with $n_k \ge N$. Taking $m = n_k$, we see that $\|f_n - f\|_p \le \|f_n - f_{n_k}\|_p + \|f_{n_k} - f\|_p \le 2\epsilon$ whenever $n \ge n_k$. Thus $\{f_n\}$ converges to f. This completes the proof of the theorem.

In Section 9 we introduced integration of functions with values in $\mathbb{R}^m$ or $\mathbb{C}^m$. The definitions of L^1, L^2, and L^∞ may be extended to include such functions, and we write $L^1(X, \mathbb{C}^m)$, for example, to indicate that the functions in question take values in $\mathbb{C}^m$. In the definitions any expression $|f(x)|$ or $|f|$ that arises in the definition and refers to absolute value in the scalar-valued case is now to be understood as referring to the norm on the vector space where the functions take their values. The vector-valued L^1, L^2, and L^∞ spaces are further normed linear spaces, and one readily checks that Theorem 5.59 with the above proof applies to them because the range spaces are complete.

The triangle inequality for a pseudo normed linear space says that the norm of the sum of two elements is less than or equal to the sum of the norms, and of course the inequality instantly extends to a sum of any finite number of elements. But what about an integral of elements? In the case that the linear space is one of the precursor spaces "V" for L^1, L^2, or L^∞, the setting is that of functions of two variables. One of the variables corresponds to the measure space under study, and the other corresponds to the indexing set for the integral of the norms. Thus we could, if we wanted, force the situation into the mold of vector-valued functions whose values are in a space of functions. But it is not necessary to do so, and we do not. Here is the theorem.

Theorem 5.60 (Minkowski's inequality for integrals). Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and put $p = 1, 2$, or ∞. If f is measurable on $X \times Y$, then

$$\Big\| \int_X f(x, y)\, d\mu(x) \Big\|_{p, d\nu(y)} \le \int_X \|f(x, y)\|_{p, d\nu(y)}\, d\mu(x)$$

in the following sense: The integrand on the right side is measurable. If the integral on the right is finite, then for almost every y $[d\nu]$ the integral on the left is defined; when it is redefined to be 0 for the exceptional y's, then the formula holds.

REMARK. An extension of this theorem to values of p other than $1, 2, \infty$ will be given in Chapter IX, and that result will have the same name.

PROOF. The right side of the integral formula is unchanged if we replace f by $|f|$, and thus we may assume that $f \ge 0$ without loss of generality. If $p = 1$, then the formula for $f \ge 0$ reads

$$\textstyle\int_Y \Big[\int_X f(x, y)\, d\mu(x) \Big] d\nu(y) \overset{?}{\le} \int_X \Big[\int_Y f(x, y)\, d\nu(y) \Big] d\mu(x).$$

In fact, equality holds, and the result just amounts to Fubini's Theorem (Theorem 5.47).

Let $p = 2$. We have
$$\|f(x, y)\|^2_{2,d\nu(y)} = \int_Y |f(x, y)|^2 \, d\nu(y),$$
and this is measurable by Fubini's Theorem. Hence $\|f(x, y)\|_{2,d\nu(y)}$ is measurable. The idea for proving the inequality in the statement of the theorem is to imitate the argument that derives the triangle inequality for L^2 from the Schwarz inequality. That earlier argument is

$$\|g + h\|_2^2 = \|g\|_2^2 + 2\,\mathrm{Re}(g, h) + \|h\|_2^2 \le \|g\|_2^2 + 2\|g\|_2\|h\|_2 + \|h\|_2^2.$$

The adapted argument is

$$\begin{aligned}\Big\| \int_X f(x, y)\, d\mu(x)\Big\|^2_{2,d\nu(y)} &= \int_Y \int_{x\in X} f(x, y)\, d\mu(x) \int_{x'\in X} \overline{f(x', y)}\, d\mu(x')\, d\nu(y)\\ &= \int_{X\times X} \Big[\int_Y f(x, y)\overline{f(x', y)} d\nu(y)\Big]\, d\mu(x)\, d\mu(x')\\ &\le \int_{X\times X} \|f(x, y)\|_{2,d\nu(y)} \|f(x', y)\|_{2,d\nu(y)}\, d\mu(x)\, d\mu(x')\\ &= \Big[\int_X \|f(x, y)\|_{2,d\nu(y)}\, d\mu(x)\Big]^2,\end{aligned}$$

the second and third lines following from Fubini's Theorem and the Schwarz inequality.

Let $p = \infty$. This is the hard case of the proof. We proceed in three steps. The first step is to prove the asserted measurability of $\|f(x, y)\|_{\infty,d\nu(y)}$, and we do so by first handling simple functions and then passing to the limit. If $s = \sum_{n=1}^N c_n I_{E_n}$ is the canonical expansion of a simple function $s \ge 0$ on $X \times Y$ and if x is fixed, then $\|s(x, y)\|_{\infty,d\nu(y)} = \max\big\{c_n \bigm| \nu((E_n)_x) > 0\big\}$. In other words, if k_n is the indicator function of the set $\big\{x \in X \bigm| \nu((E_n)_x) > 0\big\}$, then $s = \max\{c_1k_1, \dots, c_Nk_N\}$. Each function c_nk_n is measurable by Lemma 5.44, and the pointwise maximum s is measurable by Corollary 5.9. Returning to our function $f \ge 0$, we use Proposition 5.11 to choose an increasing sequence $\{s_n\}$ of nonnegative simple functions with pointwise limit f. We prove that $\|s_n(x, y)\|_{\infty,d\nu(y)}$ increases to $\|f(x, y)\|_{\infty,d\nu(y)}$ for each x, and then the measurability follows from Corollary 5.10. Since x is fixed in this step, let us drop it and consider an increasing sequence $\{s_n\}$ of nonnegative measurable functions on Y with limit f on Y; we are to show that $\|f\|_\infty = \lim \|s_n\|_\infty$. The numbers $\|s_n\|_\infty$ are monotone increasing and are $\le \|f\|_\infty$. Thus $\lim \|s_n\|_\infty \le \|f\|_\infty$. Arguing by contradiction, suppose that equality fails and that $\lim \|s_n\|_\infty \le M < M+\epsilon < \|f\|_\infty$. Then $\big\{y \bigm| s_n(y) \ge M+\epsilon\big\}$ has measure 0 for every n, and so does $\bigcup_n \big\{y \bigm| s_n(y) \ge M + \epsilon\big\}$, by complete additivity. On the other hand, $\big\{y \bigm| f(y) > M + \epsilon\big\}$ is a subset of this union, and it has positive measure since $M + \epsilon < \|f\|_\infty$. Thus we have a contradiction and conclude that $\lim \|s_n\|_\infty = \|f\|_\infty$. Consequently $\|f(x, y)\|_{\infty,d\nu(y)}$ is measurable, as asserted.

The second step is to prove that any measurable function $F \geq 0$ on Y has $\|F\|_\infty = \sup_g \left| \int_Y Fg\,d\nu \right|$, where the supremum is taken over all $g \geq 0$ with $\|g\|_1 \leq 1$. Certainly any such g has $\left| \int_Y Fg\,d\nu \right| \leq \|F\|_\infty \int_Y g\,d\nu = \|F\|_\infty$, and therefore $\sup_g \left| \int_Y Fg\,d\nu \right| \leq \|F\|_\infty$. For the reverse inequality, let I_E be the indicator function of a set of finite positive measure, and put $g = \nu(E)^{-1} I_E$. Then $\int_Y Fg\,d\nu = \nu(E)^{-1} \int_E F\,d\nu \geq \inf_E(F)$. If m is less than $\|F\|_\infty$, then the set E where F is $\geq m$ has positive measure, and the inequality reads $m \leq \int_Y Fg\,d\nu$ for the associated g. Hence $m \leq \sup_g \int_Y Fg\,d\nu$. Taking the supremum of such m's, we obtain $\|F\|_\infty \leq \sup_g \left| \int_Y Fg\,d\nu \right|$, and the reverse inequality is proved.

The third step is to use the previous two steps to prove the inequality in the statement of the theorem for $f \geq 0$. Let g be any nonnegative function on Y with $\int_Y g\,d\nu \leq 1$. Then Fubini's Theorem, the result of the first step above, and the result in the easy direction of the second step above give

$$\begin{aligned}\int_Y g(y)\Big[\int_X f(x,y)\,d\mu(x)\Big]\,d\nu(y) &= \int_X \Big[\int_Y f(x,y)g(y)\,d\nu(y)\Big]\,d\mu(x)\\ &\leq \int_X \Big[\|f(x,y)\|_{\infty,d\nu(y)}\Big]\,d\mu(x).\end{aligned}$$

Taking the supremum over g and using the result in the hard direction of the second step, we obtain the inequality in the statement of the theorem.

10. Problems

1. Let X be a finite set of $n > 0$ elements.
   (a) If $\mathcal{A}$ is an algebra of subsets, what are the possible numbers of sets in $\mathcal{A}$?
   (b) Show that symmetric difference $A \,\Delta\, B = (A - B) \cup (B - A)$ is an abelian group operation on the set of all subsets of X and that every nontrivial element has order 2.
   (c) If $\mathcal{B}$ is a class of subsets containing $\varnothing$ and X and closed under symmetric difference, what are the possible numbers of sets in $\mathcal{B}$?
   (d) Prove or disprove: The class of sets in (c) is necessarily an algebra of sets.
   (e) Show that intersection and symmetric difference satisfy the distributive law $A \cap (B \,\Delta\, C) = (A \cap B) \,\Delta\, (A \cap C)$.
2. Exhibit a completely additive set function ρ on a σ-algebra and two sets A and B such that $\rho(A) < 0$ and $\rho(B) < 0$ but $\rho(A \cup B) > 0$.
3. Let $\{E_n\}$ be a sequence of subsets of X, and put

$$A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k \quad \text{and} \quad B = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k.$$

Prove that the indicator functions of E_k, A, and B satisfy

$$I_A = \limsup_n I_{E_n} \quad \text{and} \quad I_B = \liminf_n I_{E_n}.$$

4. Suppose that μ is a finite measure defined on a σ-algebra and $\{E_n\}$ is a sequence of measurable sets with

$$\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k.$$

Call the set on the two sides of this equation E. Prove that $\lim_n \mu(E_n)$ exists and equals $\mu(E)$.

5. Let X be the set of rational numbers, and let $\mathcal{R}$ be the ring of all finite disjoint unions of bounded intervals in X, with or without endpoints. For each set E in $\mathcal{R}$, let $\mu(E)$ be its length.
   (a) Show that μ is nonnegative additive.
   (b) Show that μ is not completely additive.
6. Prove that if E is a Lebesgue measurable subset of [0, 1] of Lebesgue measure 0, then the complement of E is dense in [0, 1].
7. Let μ be a measure defined on a σ-algebra. Prove that if the complement of every set of measure $+\infty$ is of finite measure, then $\sup_{\mu(A)<+\infty} \mu(A)$ is finite and there is a set B with $\mu(B) = \sup_{\mu(A)<+\infty} \mu(A)$.
8. If f is a measurable function, prove that $f^{-1}(E)$ is measurable whenever E is a Borel subset of the real line.
9. For the measure space $(X, \mathcal{A}, \mu)$ in which X is the positive integers, $\mathcal{A}$ consists of all subsets of X, and μ is the counting measure, the theory of Lebesgue integration becomes a theory of infinite series. Restate Fatou's Lemma and the Dominated Convergence Theorem in this context.
10. Suppose on a finite measure space that $\{f_n\}$ is a sequence of real-valued integrable functions tending uniformly to f. Prove that $\lim_n \int_X f_n\,d\mu = \int_X f\,d\mu$.
11. This problem involves a Cantor set C in [0, 1] built using fractions r_n as in Section II.9.
    (a) Show that C has Lebesgue measure $\prod_{n=1}^{\infty} (1 - r_n)$.
    (b) Prove that the indicator function I_C is discontinuous at every point of C and only there. Thus the set of discontinuities of I_C is not of measure 0 if $\prod_{n=1}^{\infty} (1 - r_n) > 0$.
    (c) Show that if the result of redefining I_C on a set of Lebesgue measure 0 is a function f, then the only possible points of continuity of f are those where f is 0.
    (d) Conclude that there exists a Lebesgue measurable function on [0, 1] that is not Riemann integrable and cannot be redefined on a set of measure 0 so as to be Riemann integrable.

12. Let $(X, \mathcal{A}, \mu)$ be any measure space, and let $(X, \overline{\mathcal{A}}, \overline{\mu})$ be its completion. Prove that if f is a function measurable with respect to $\overline{\mathcal{A}}$, then f can be redefined on a set of $\overline{\mu}$-measure 0 so as to be measurable with respect to $\mathcal{A}$.

13. Let X be an uncountable set, and let $\mathcal{A}$ be the set of all countable subsets of X and their complements. Prove that the diagonal $\{(x, x) \mid x \in X\}$ is not a member of the σ-algebra $\mathcal{A} \times \mathcal{A}$, the smallest σ-algebra containing all rectangles with sides in $\mathcal{A}$.

14. Let $(\mathbb{R}^1, \mathcal{B}, m)$ be the real line with Lebesgue measure on the Borel sets, and let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. If $f \geq 0$ is a measurable function on X, prove that the "region under the graph of f," defined by
$$R = \{(x, y) \mid 0 \leq y < f(x)\},$$
is a measurable subset of $X \times \mathbb{R}^1$ and that its measure relative to $\mu \times m$ is $\int_X f(x)\, d\mu(x)$.

15. Let $\mathcal{A}$ be a σ-algebra of subsets of a nonempty set X, let $F : \mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_k} \to \mathbb{C}^N$ be continuous, and let $f_j : X \to \mathbb{C}^{n_j}$ be measurable with respect to $\mathcal{A}$ for $1 \leq j \leq k$. Prove that $x \mapsto F(f_1(x), \ldots, f_k(x))$ is measurable with respect to $\mathcal{A}$.

16. This problem complements the proof in Theorem 5.59 that L^1 is a complete metric space. For $n \geq 1$, suppose that $0 < a_n < 1$ and $\sum_{n=1}^{\infty} a_n = +\infty$. Find a measure space $(X, \mathcal{A}, \mu)$ and a sequence of functions f_n with $\|f_n\|_1 = a_n$ and $\{f_n(x)\}$ convergent for no x.

17. **(Egoroff's Theorem)** Let $(X, \mathcal{A}, \mu)$ be a finite measure space. Suppose that f_n and f are measurable functions with values in $\mathbb{R}$ such that $\lim f_n(x) = f(x)$ pointwise. The objective of this problem is to prove that $\lim f_n = f$ "almost uniformly." By considering the sets
$$E_{MN} = \{x \in X \mid |f_n(x) - f(x)| < 1/M \text{ for } n \geq N\}$$
for M fixed and N varying, prove that if $\epsilon > 0$ is given, then there exists a measurable subset E of X with $\mu(E) < \epsilon$ such that $\lim f_n(x) = f(x)$ uniformly for x in E^c.

18. (a) Derive the Dominated Convergence Theorem for a space of finite measure from Egoroff's Theorem (Problem 17) and Corollary 5.24.
    (b) Derive the Dominated Convergence Theorem for a space of infinite measure from the Dominated Convergence Theorem for a space of finite measure.

Problems 19–21 use Egoroff's Theorem (Problem 17) to show how close pointwise convergence is to L^1 convergence on a measure space $(X, \mathcal{A}, \mu)$ of finite measure. Theorem 5.59 shows that if a sequence converges in $L^1(X)$, then a subsequence converges almost everywhere. These problems address the converse direction in a way different from Problem 16. Suppose that f_n and f are integrable functions with values in $\mathbb{R}$ such that $\lim f_n(x) = f(x)$ pointwise.

19. Suppose that $f_n \geq 0$ for all n and that $\lim \int_X f_n \, d\mu = \int_X f \, d\mu$. Prove that $\lim_n \int_E f_n \, d\mu = \int_E f \, d\mu$ for every measurable set E.

20. Suppose that $f_n \geq 0$ for all n and that $\lim \int_X f_n \, d\mu = \int_X f \, d\mu$. Use the previous problem and Egoroff's Theorem to prove that $\lim \int_X |f_n - f| \, d\mu = 0$.

21. A sequence $\{g_n\}$ of nonnegative integrable functions is called **uniformly integrable** if for any $\epsilon > 0$, there is an N such that $\int_{\{x \mid f_n(x) \geq N\}} g_n \, d\mu < \epsilon$ for all n. Suppose that the members of the given convergent sequence $\{f_n\}$ are nonnegative. Using Egoroff's Theorem in one direction and the previous problem in the converse direction, prove that $\lim_n \int_X f_n \, d\mu = \int_X f \, d\mu$ if and only if the f_n are uniformly integrable.

Problems 22–24 concern the extension of measures beyond what is given in Theorem 5.5 and Proposition 5.37. Let μ be a finite measure on a σ-algebra $\mathcal{A}$ of subsets of X, and define μ_* and μ^* on all subsets of X as in Lemma 5.32 and immediately after it. Let E be a subset of X that is not in $\mathcal{A}$, and let $\mathcal{B}$ be the smallest σ-algebra containing E and the members of $\mathcal{A}$.

22. Show that there exist two sets K and U in $\mathcal{A}$ such that $K \subseteq E \subseteq U$, $\mu_*(E) = \mu(K)$, and $\mu^*(E) = \mu(U)$. Show that K and U have the further properties that $U^c \subseteq E^c \subseteq K^c$, $\mu_*(E^c) = \mu(U^c)$, and $\mu^*(E^c) = \mu(K^c)$.

23. Show that the sets K and U of the previous problem satisfy $\mu_*(A \cap E) = \mu(A \cap K)$ and $\mu^*(A \cap E) = \mu(A \cap U)$ for every A in $\mathcal{A}$.

24. Fix t in [0, 1]. Show that the set function σ defined for A and B in $\mathcal{A}$ by
$$\begin{aligned}&\sigma[(A \cap E) \cup (B \cap E^c)] \\ &\qquad = t\mu_*(A \cap E) + (1-t)\mu^*(A \cap E) + t\mu^*(B \cap E^c) + (1-t)\mu_*(B \cap E^c)\end{aligned}$$
is defined on all of $\mathcal{B}$, is a measure, agrees with μ on $\mathcal{A}$, and assigns measure $t\mu_*(E) + (1-t)\mu^*(E)$ to the set E.

Problems 25–33 concern a construction by "transfinite induction" of all sets in the smallest σ-algebra containing an algebra of sets. In particular, it describes how to obtain all Borel sets of the interval [0, 1] of the line from the elementary sets in that interval. Later problems in the set apply the construction in various ways. This set of problems makes use of partial orderings as described in Section A9 of the appendix, but they do not use Zorn's Lemma. The set of **countable ordinals** is an uncountable partially ordered set Ω, under a partial ordering $\leq$, with the following properties:

(i) Ω has the property that $x \leq y$ and $y \leq x$ together imply $x = y$,
(ii) Ω is "totally ordered" in the sense that any x and y in the set have either $x \leq y$ or $y \leq x$,
(iii) Ω is "well ordered" in the sense that any nonempty subset has a least element,
(iv) for any x in Ω, the set of elements $\leq x$ is at most countable.

Take as known that such a set Ω exists.

25. Prove that any countable subset of Ω has a least upper bound.

26. This problem asks for a proof of the validity of **transfinite induction** as applied to Ω. Let 1 be the least element of Ω, and let "$<$" mean "$\le$ but not $=$." Suppose that some $p(\omega)$ is specified for each ω in Ω. Suppose further that $p(1)$ is true and that if for each $\omega > 1$, $p(\omega')$ is true for all $\omega' < \omega$, then $p(\omega)$ is true. Prove that $p(\omega)$ is true for all ω in Ω.

27. Let X be a nonempty set, let $\mathcal{A}$ be an algebra of subsets of X, and let $\mathcal{B}$ be the smallest σ-algebra containing $\mathcal{A}$. This problem uses Ω to describe "constructively" $\mathcal{B}$ in terms of $\mathcal{A}$. We define by transfinite induction two successively larger classes of sets $\mathcal{U}_\alpha$ and $\mathcal{K}_\alpha$ for each countable ordinal $\alpha \ge 1$. Let $\mathcal{U}_1$ be the set of all countable increasing unions of members of $\mathcal{A}$, let $\mathcal{K}_\alpha$ for $\alpha \ge 1$ be the set of all countable decreasing intersections of members of $\mathcal{U}_\alpha$, and let $\mathcal{U}_\alpha$ for $\alpha > 1$ be the set of all countable increasing unions of members of previous $\mathcal{K}_\beta$'s.
   (a) Prove at each stage α that $\mathcal{U}_\alpha$ and $\mathcal{K}_\alpha$ are both closed under finite unions and finite intersections.
   (b) Prove that $\mathcal{B}$ is the union of all $\mathcal{K}_\alpha$ for α in Ω.

28. For the case that $\nu(X) < +\infty$, prove the uniqueness half of the Extension Theorem (Theorem 5.5) by using the transfinite construction of Problem 27. [Educational note: It is not known how to prove the existence half of the Extension Theorem in this "constructive" way.]

29. Prove the Monotone Class Lemma (Lemma 5.43) by making use of the transfinite construction of Problem 27.

30. Devise a transfinite construction of all finite-valued Borel measurable functions on $\mathbb{R}^1$ that starts from continuous functions and alternately allows pointwise increasing limits and pointwise decreasing limits. The construction is to be in the spirit of Problem 27. Show that all finite-valued Borel measurable functions are obtained in this way if the indexing is done with Ω.

31. This problem "counts" the number of Borel sets of the real line, using Problem 27. It uses the material on cardinality in Section A10 of the appendix.
   (a) Prove that
      (i) Ω has the same cardinality as some subset of $\mathbb{R}$,
      (ii) the set of all sequences of members of $\mathbb{R}$ has the same cardinality as $\mathbb{R}$,
      (iii) if $A \subseteq B \subseteq C$ and if A and C have the same cardinality as $\mathbb{R}$, then so does B,
      (iv) if a set A has the same cardinality as $\mathbb{R}$ and if for each α in A, B_α is a set with the same cardinality as $\mathbb{R}$, then $\bigcup_{\alpha\in A} B_\alpha$ has the same cardinality as $\mathbb{R}$.
   (b) Deduce that the set of all Borel sets of $\mathbb{R}$ has the same cardinality as $\mathbb{R}$ itself.

32. The standard Cantor set C in $[0, 1]$, built using fractions $r_n = 1/3$ as in Section II.9, is a Borel set of Lebesgue measure 0 by Problem 11. Prove that C has the same cardinality as $\mathbb{R}$. Conclude that the cardinality of the set of all Lebesgue measurable sets equals the cardinality of the set of all subsets of $\mathbb{R}$. [Educational note: From this and Problem 31 it follows that there exists a Lebesgue measurable set in $[0, 1]$ that is not a Borel set.]

33. For the standard Cantor set C as in the previous problem, show that the indicator function $I_{C'}$ of any subset C' of C is continuous on C^c. Conclude that the cardinality of the set of Riemann integrable functions on $[0, 1]$ equals the cardinality of the set of all subsets of $\mathbb{R}$. [Educational note: From this and Problems 30–31, it follows that there exists a Riemann integrable function on $[0, 1]$ that is not Borel measurable.]

Problems 34–41 show how to produce nontrivial nonnegative additive set functions on the set of all subsets of an infinite set from Zorn's Lemma (Section A9 of the appendix). A **filter** $\mathcal{F}$ on a nonempty set X is a nonempty class of subsets of X such that

(i) if E is in $\mathcal{F}$ and $F \supseteq E$, then F is in $\mathcal{F}$, i.e., $\mathcal{F}$ is closed under the operation of forming supersets,

(ii) if E and F are in $\mathcal{F}$, so is $E \cap F$,

(iii) $\varnothing$ is not in $\mathcal{F}$.

An **ultrafilter** is a filter that is not properly contained in any larger filter.

34. Verify the following:
    (a) $\{X\}$ is a filter.
    (b) Any filter is closed under finite intersections.
    (c) A one-point set and all of its supersets form an ultrafilter. (Such an ultrafilter is called a **trivial ultrafilter**.)
    (d) If X is infinite, then the set $\mathcal{F}$ of all subsets whose complements are finite sets is a filter.

35. Use Zorn's Lemma to show that every filter is contained in some ultrafilter.

36. Show that if $\mathcal{C}$ is a nonempty class of subsets of X, then there is a filter containing $\mathcal{C}$ if and only if no finite intersection of members of $\mathcal{C}$ is empty.

37. Prove that a filter $\mathcal{F}$ is an ultrafilter if and only if $A \cup B$ in $\mathcal{F}$ implies that either A is in $\mathcal{F}$ or B is in $\mathcal{F}$.

38. Prove that a filter $\mathcal{F}$ is an ultrafilter if and only if for every $A \subseteq X$, either A is in $\mathcal{F}$ or A^c is in $\mathcal{F}$.

39. Prove that the nonzero additive set functions defined on the set of *all* subsets of a set X and having image $\{0, 1\}$ stand in one-one correspondence with the ultrafilters on X, the correspondence being that the sets in the ultrafilter are exactly the sets on which the set function is 1. Prove that the set function is a measure if and only if the corresponding ultrafilter is closed under countable intersections.

40. Let X be any infinite set. Prove that X has a nontrivial ultrafilter, hence that X has a nonnegative additive set function μ that assumes only the values 0 and 1 and is not a point mass.

41. Prove that the set $\mathbb{Z}^+$ of positive integers has no nontrivial ultrafilter closed under countable intersections, i.e., that the set function μ in the previous problem is not a measure.

Problems 42–43 concern a theory of integration in which complete additivity is dropped as an assumption. An example is given in Problems 39–41 of a nonnegative additive set function on the set of all subsets of an infinite set that is not completely additive. For the present set of problems, let X be a nonempty set, let $\mathcal{A}$ be a σ-algebra of subsets, and let μ be a nonnegative additive set function on $\mathcal{A}$ such that $\mu(X) < +\infty$. Imagine an integration theory for $\int_E f\,d\mu$ with the definitions just as in the case that μ is a measure. All the properties of the integral proved in the text before the Monotone Convergence Theorem would still be valid, except that the integral $\int_E f\,d\mu$ as a function of E would be merely additive, rather than completely additive, and hence we would have to drop Corollary 5.24 and the converse half of Corollary 5.23.

42. Let f be ≥ 0, and let s_n be the standard pointwise increasing sequence of simple functions with limit f, as in Proposition 5.11. Show that the convergence of s_n to f is uniform if f is bounded.

43. Use the result of the previous problem to show in this theory that $\int_E (f+g)\,d\mu = \int_E f\,d\mu + \int_E g\,d\mu$ if f and g are bounded and measurable.

CHAPTER VI

Measure Theory for Euclidean Space

Abstract. This chapter mines some of the powerful consequences of the basic measure theory in Chapter V.

Sections 1–3 establish properties of Lebesgue measure and other Borel measures on Euclidean space and on open subsets of Euclidean space. The main general property is the regularity of all such measures—that the measure of any Borel set can be approximated by the measure of compact sets from within and open sets from without. Lebesgue measure in all of Euclidean space has an additional property, translation invariance, which allows for the notion of the convolution of two functions. Convolution gives a kind of moving average of the translates of one function weighted by the other function. Convolution with the dilates of a fixed integrable function provides a handy kind of approximate identity.

Section 4 gives the final form of the comparison of the Riemann and Lebesgue integrals, a preliminary form having been given in Chapter III.

Section 5 gives the final form of the change-of-variables theorem for integration, starting from the preliminary form of the theorem in Chapter III and taking advantage of the ease with which limits can be handled by the Lebesgue integral. Sard's Theorem allows one to disregard sets of lower dimension in establishing such changes of variables, thereby giving results in their expected form rather than in a form dictated by technicalities.

Section 6 concerns the Hardy–Littlewood Maximal Theorem in N dimensions. In dimension 1, this theorem implies that the derivative of a 1-dimensional Lebesgue integral with respect to Lebesgue measure recovers the integrand almost everywhere. The theorem in the general case implies that certain averages of a function over small sets about a point tend to the function almost everywhere. But the theorem can be regarded as saying also that a particular approximate identity formed by dilations applies to problems of almost-everywhere convergence, as well as to problems of norm convergence and uniform convergence. A corollary of the theorem is that many approximate identities formed by dilations yield almost-everywhere convergence theorems.

Section 7 redevelops the beginnings of the subject of Fourier series using the Lebesgue integral, the theory having been developed with the Riemann integral in Section I.10. With the Lebesgue integral and its accompanying tools, Fourier series are meaningful for more functions than before, Dini's test applies even to a wider class of Riemann integrable functions than before, and Fejér's Theorem and Parseval's Theorem become easier and more general than before. A completely new result with the Lebesgue integral is the Riesz–Fischer Theorem, which characterizes the trigonometric series that are Fourier series of square-integrable functions.

Sections 8–10 deal with Stieltjes measures, which are Borel measures on the line, and their application to Fourier series. Such measures are characterized in terms of a class of monotone functions on the line, and they lead to a handy generalization of the integration-by-parts formula. This formula allows one to bound the size of the Fourier coefficients of functions of bounded variation, which are differences of monotone functions. In combination with earlier results, this bound yields

the Dirichlet–Jordan Theorem, which says that the Fourier series of a function of bounded variation converges pointwise everywhere, the convergence being uniform on any compact set on which the function is continuous. Section 10 is a short section on computation of integrals.

1. Lebesgue Measure and Other Borel Measures

Lebesgue measure on $\mathbb{R}^1$ was constructed in Section V.1 on the ring of "elementary" sets—the finite disjoint unions of bounded intervals—and extended from there to the σ-algebra of Borel sets by the Extension Theorem (Theorem 5.5), which was proved in Section V.5. Fubini's Theorem (Theorem 5.47) would have allowed us to build Lebesgue measure in $\mathbb{R}^N$ as an iterated product of 1-dimensional Lebesgue measure, but we postponed the construction in $\mathbb{R}^N$ until the present chapter in order to show that it can be carried out in a fashion independent of how we group 1-dimensional factors.

The Borel sets of $\mathbb{R}^1$ are, by definition, the sets in the smallest σ-algebra containing the elementary sets, and we saw readily that every set that is open or compact is a Borel set. We write $\mathcal{B}_1$ for this σ-algebra. In fact, $\mathcal{B}_1$ may be described as the smallest σ-algebra containing the open sets of $\mathbb{R}^1$ or as the smallest σ-algebra containing the compact sets. The reason that the open sets generate $\mathcal{B}_1$ is that every open interval is an open set, and every interval is a countable intersection of open intervals. Similarly the compact sets generate $\mathcal{B}_1$ because every closed bounded interval is a compact set, and every interval is the countable union of closed bounded intervals.

Now let us turn our attention to $\mathbb{R}^N$. We have already used the word "rectangle" in two different senses in connection with integration—in Chapter III to mean an N-fold product along coordinate directions of open or closed bounded intervals, and in Chapter V to mean a product of measurable sets. For clarity let us refer to any product of bounded intervals as a **geometric rectangle** and to any product of measurable sets as an **abstract rectangle** or an abstract rectangle in the sense of Fubini's Theorem. In $\mathbb{R}^N$, every geometric rectangle under our definition is an abstract rectangle, but not conversely.

Define the **Borel sets** of $\mathbb{R}^N$ to be the members of the smallest σ-algebra $\mathcal{B}_N$ containing all compact sets in $\mathbb{R}^N$. It is equivalent to let $\mathcal{B}_N$ be the smallest σ-algebra containing all open sets. In fact, every open geometric rectangle is the countable union of compact geometric rectangles, and every open set in turn is the countable union of open geometric rectangles; thus the open sets are in the smallest σ-algebra containing the compact sets. In the reverse direction every closed set is the complement of an open set, and every compact set is closed; thus the compact sets are in the smallest σ-algebra containing the open sets.

Functions on $\mathbb{R}^N$ measurable with respect to $\mathcal{B}_N$ are called **Borel measurable functions** or **Borel functions**. Any continuous real-valued function f on $\mathbb{R}^N$

is Borel measurable because the inverse image $f^{-1}((c, +\infty])$ of the open set $(c, +\infty]$ has to be open and therefore has to be a Borel set.

Proposition 6.1. If m and n are integers ≥ 1, then $\mathcal{B}_m \times \mathcal{B}_n = \mathcal{B}_{m+n}$ within the product set $\mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$.

PROOF. If U is open is $\mathbb{R}^m$ and V is open in $\mathbb{R}^n$, then $U \times V$ is open in $\mathbb{R}^{m+n}$, and it follows that $\mathcal{B}_m \times \mathcal{B}_n \subseteq \mathcal{B}_{m+n}$. For the reverse inclusion, let W be open in $\mathbb{R}^{m+n}$. Then W is the countable union of open geometric rectangles, and each of these is of the form $U \times V$ with U open in $\mathbb{R}^m$ and V open in $\mathbb{R}^n$. Since each such $U \times V$ is in $\mathcal{B}_m \times \mathcal{B}_n$, so is W. Thus we obtain the reverse inclusion $\mathcal{B}_{m+n} \subseteq \mathcal{B}_m \times \mathcal{B}_n$.

Lebesgue measure on $\mathbb{R}^N$ will, at least initially, be a measure defined on the σ-algebra $\mathcal{B}_N$. Proposition 6.1 tells us that the σ-algebra on which the measure is to be defined is independent of the grouping of variables used in Fubini's Theorem. It will be quite believable that different constructions of Lebesgue measure by using different iterated product decompositions of $\mathbb{R}^N$, such as $(\mathbb{R}^1 \times \mathbb{R}^1) \times \mathbb{R}^1$ and $\mathbb{R}^1 \times (\mathbb{R}^1 \times \mathbb{R}^1)$, will lead to the same measure, but we shall give two abstract characterizations of the result that will ensure uniqueness without any act of faith. These characterizations will take some moments to establish, but we shall obtain useful additional results along the way. The procedure will be to state the constructions of the measure via Fubini's Theorem, then to consider a wider class of measures on $\mathcal{B}_N$ known as the "Borel measures," and finally to establish the two characterizations of Lebesgue measure among all Borel measures on $\mathbb{R}^N$.

It is customary to write dx in place of $dm(x)$ for Lebesgue measure on $\mathbb{R}^1$, and we shall do so except when there is some special need for the symbol m. Then the notation for the measure normally becomes an expression like dx or dy instead of m. To construct Lebesgue measure dx on $\mathbb{R}^N$, we can proceed inductively, adding one variable at a time. Fubini's Theorem allows us to construct the product of Lebesgue measure on $\mathbb{R}^{N-1}$ and Lebesgue measure on $\mathbb{R}^1$, and Proposition 6.1 shows that the result is defined on the Borel sets of $\mathbb{R}^N$. Let us take this particular construction as an inductive definition of **Lebesgue measure** on $\mathbb{R}^N$. It is apparent from the construction that the measure of a geometric rectangle is the product of the lengths of the sides.

Alternatively, we could construct Lebesgue measure on $\mathbb{R}^N$ inductively by grouping $\mathbb{R}^N$ as some other $\mathbb{R}^m \times \mathbb{R}^{N-m}$ and using the product measure from versions of the Lebesgue measures on $\mathbb{R}^m$ and $\mathbb{R}^{N-m}$. Again the result has the property that the measure of a geometric rectangle is the product of the lengths of the sides. It is believable that this condition determines completely the measure on $\mathbb{R}^N$, and we shall give a proof of this uniqueness shortly.

A **Borel measure** on $\mathbb{R}^N$ is a measure on the σ-algebra $\mathcal{B}_N$ of Borel sets of $\mathbb{R}^N$ that is finite on every compact set. A key property of Borel measures on $\mathbb{R}^N$ is their regularity as expressed in Theorem 6.2 below. The theorem makes use of two simple properties of $\mathbb{R}^N$:

(i) there exists a sequence $\{F_n\}_{n=1}^{\infty}$ of compact sets with union the whole space such that $F_n \subseteq F_{n+1}^o$ for all n,

(ii) for any compact set K, there exists a decreasing sequence of open sets U_n with compact closure such that $\bigcap_{n=1}^{\infty} U_n = K$.

For (i), we can take F_n to be the closed ball of radius n centered at the origin. For (ii), we can take $U_n = \{x \mid D(x, K) < 1/n\}$ if $K \neq \varnothing$, and we can take all $U_n = \varnothing$ if $K = \varnothing$.

Theorem 6.2. Every Borel measure μ on $\mathbb{R}^N$ is **regular** in the sense that the value of μ on any Borel set E is given by

$$\mu(E) = \sup_{\substack{K \subseteq E,\\ K \text{ compact}}} \mu(K) = \inf_{\substack{U \supseteq E,\\ U \text{ open}}} \mu(U).$$

REMARK. This conclusion is new for us even for $\mathbb{R}^1$. Although regularity of 1-dimensional Lebesgue measure was introduced before Proposition 5.4, it was established only for the elementary sets at that time.

PROOF. We shall begin by showing for each Borel set E and for any $\epsilon > 0$ that

$$\begin{aligned}&\text{there exist closed } C \text{ and open } U \text{ such that}\\ &C \subseteq E \subseteq U \text{ and } \mu(U - C) < \epsilon.\end{aligned} \tag{$*$}$$

Let $\mathcal{A}$ be the set of Borel sets E for which $(*)$ holds for all $\epsilon > 0$.

If E is compact, then we can take $C = E$ and $U = U_n$ as in (ii) for a suitable n in order to prove $(*)$; Corollary 5.3 gives us $\lim_n \mu(U_n - C) = 0$, since the compact closure of U_n forces $\mu(U_1)$ to be finite. Therefore $\mathcal{A}$ contains all compact sets.

To see that $\mathcal{A}$ is closed under complements, suppose E is in $\mathcal{A}$. Let $\epsilon > 0$ be given and choose, by $(*)$ for E, a closed set C and an open set U such that $C \subseteq E \subseteq U$ and $\mu(U - C) < \epsilon$. Taking complements, we have $U^c \subseteq E^c \subseteq C^c$ and $\mu(C^c - U^c) = \mu(U - C) < \epsilon$. Thus E^c is in $\mathcal{A}$.

Let us see that $\mathcal{A}$ is closed under finite unions. Suppose that E_1 and E_2 are in $\mathcal{A}$. Let $\epsilon > 0$ be given and choose, by $(*)$ for E_1 and E_2, two closed sets C_1 and C_2 and two open sets U_1 and U_2 such that $C_1 \subseteq E_1 \subseteq U_1$, $\mu(U_1 - C_1) < \epsilon$, $C_2 \subseteq E_2 \subseteq U_2$, and $\mu(U_2 - C_2) < \epsilon$. Then $C_1 \cup C_2 \subseteq E_1 \cup E_2 \subseteq U_1 \cup U_2$ and $\mu((U_1 \cup U_2) - (C_1 \cup C_2)) \leq \mu(U_1 - C_1) + \mu(U_2 - C_2) < 2\epsilon$. Since ϵ is arbitrary, $E_1 \cup E_2$ is in $\mathcal{A}$. Hence $\mathcal{A}$ is closed under finite unions, and $\mathcal{A}$ is an algebra of sets.

The proof that $\mathcal{A}$ is closed under countable unions takes two steps. For the first step we let a sequence of sets E_n in $\mathcal{A}$ be given with union E, and first assume that all E_n lie in one of the sets F_M in (i) above. Let $\epsilon > 0$ be given and choose, by $(*)$ for each E_n, closed sets C_n and open sets U_n such that $C_n \subseteq E_n \subseteq U_n$ and $\mu(U_n - C_n) < \epsilon/2^n$. Possibly by intersecting U_n with F^o_{M+1}, we may assume that all U_n lie in the compact set F_{M+1}. Set $U = \bigcup_{n=1}^\infty U_n$ and $C = \bigcup_{n=1}^\infty C_n$. Then $C \subseteq E \subseteq U$ with U open but C not necessarily closed. Nevertheless, we have $U - C \subseteq \bigcup_{n=1}^\infty (U_n - C_n)$, and Proposition 5.1g gives $\mu(U - C) \le \sum_{n=1}^\infty \mu(U_n - C_n) < \epsilon$. The sets $S_m = U - \bigcup_{n=1}^m C_n$ form a decreasing sequence within F_{M+1} with intersection $U - C$. Since $\mu(F_{M+1})$ is finite, Corollary 5.3 shows that $\mu(S_m)$ decreases to $\mu(U - C)$, which is $< \epsilon$. Thus there is some $m = m_0$ with $\mu(S_{m_0}) < \epsilon$. The set $C' = \bigcup_{n=1}^{m_0} C_n$ is closed, and we have $C' \subseteq E \subseteq U$ and $\mu(U - C') = \mu(S_{m_0}) < \epsilon$. Therefore E is in $\mathcal{A}$.

For the second step we let the sets E_n be general members of $\mathcal{A}$. Since $\mathcal{A}$ is an algebra, $E_n \cap (F_{m+1} - F_m)$ is in $\mathcal{A}$ for every n and m. Applying the previous step, we see that $E'_m = E \cap (F_{m+1} - F_m)$ is in $\mathcal{A}$ for every m. The sets E'_m have union E, and E'_m is contained in $F_{m+1} - F_n$. Changing notation, we may assume that the given sets E_n all have $E_n \subseteq F_{n+1} - F_n$. If $\epsilon > 0$ is given, construct U_n open and C_n closed as in the previous paragraph except that U_n is not constrained to lie in a particular F_M. Again let $U = \bigcup_{n=1}^\infty U_n$ and $C = \bigcup_{n=1}^\infty C_n$, so that $C \subseteq E \subseteq U$ and $\mu(U - C) < \epsilon$. The set U is open, and this time we can prove that the set C is closed. In fact, let $\{x_k\}$ be a sequence in C convergent to some limit point x_0 of C. The point x_0 is in some F_M since the sets F_M have union the whole space. Since $F_M \subseteq F^o_{M+1}$ and F^o_{M+1} is open, the sequence is eventually in F^0_{M+1}. The inclusion $C_n \subseteq E_n \subseteq F_{n+1} - F_n$ shows that $C_n \cap F_{M+1} = \varnothing$ for $n \ge M + 1$. Thus no term of the sequence after some point lies in $C_{M+1}, C_{M+2}, \dots$, i.e., all the terms of the sequence after some point lie in $\bigcup_{n=1}^M C_n$. This is a closed set, and the limit x_0 must lie in it. Therefore x_0 lies in C, and C is closed. This proves that E is in $\mathcal{A}$. Hence $\mathcal{A}$ is a σ-algebra and must contain all Borel sets.

From $(*)$ for all Borel sets, it follows that every Borel set E satisfies

$$\mu(E) = \sup_{\substack{C \subseteq E, \\ C \text{ closed}}} \mu(C) = \inf_{\substack{U \supseteq E, \\ U \text{ open}}} \mu(U). \qquad (**)$$

Proposition 5.2 shows that the sets F_n of (i) have the property that $\mu(C) = \sup \mu(C \cap F_n)$ for every Borel set C. When C is closed, the sets $C \cap F_n$ are compact, and thus $(**)$ implies the equality asserted in the statement of the theorem. This completes the proof.

Recall from Section III.10 that the **support** of a scalar-valued function on a metric space is the closure of the set where it is nonzero. Let $C_{\text{com}}(\mathbb{R}^N)$ be the space of continuous scalar-valued functions on $\mathbb{R}^N$ of compact support. If there is no special mention of the scalars, the scalars may be either real or complex.

If K is a compact set and the open sets U_n are as in (ii) before Theorem 6.2, Proposition 2.30e gives us continuous functions $f_n : \mathbb{R}^N \to [0, 1]$ such that f_n is 1 on K and is 0 on U_n^c. The support of the function f_n is then contained in U_n^{cl}, which is compact. By replacing the functions f_n by $g_n = \min\{f_1, \dots, f_n\}$, we may assume that they are pointwise decreasing. Consequently

(iii) there exists a decreasing sequence of real-valued members of $C_{\text{com}}(\mathbb{R}^N)$ with pointwise limit the indicator function of K.

Corollary 6.3. If μ and ν are Borel measures on $\mathbb{R}^N$ such that $\int_{\mathbb{R}^N} f\, d\mu = \int_{\mathbb{R}^N} f\, d\nu$ for all continuous functions on $\mathbb{R}^N$ of compact support, then $\mu = \nu$.

PROOF. Let K be a compact subset of $\mathbb{R}^N$, and use (iii) to choose a decreasing sequence $\{f_n\}$ of real-valued members of $C_{\text{com}}(\mathbb{R}^N)$ with pointwise limit the indicator function I_K. Since f_1 is integrable, dominated convergence allows us to deduce $\int_{\mathbb{R}^N} I_K\, d\mu = \int_{\mathbb{R}^N} I_K\, d\nu$ from the equality $\int_{\mathbb{R}^N} f_n\, d\mu = \int_{\mathbb{R}^N} f_n\, d\nu$ for all n. Thus $\mu(K) = \nu(K)$ for every compact set K. Applying Theorem 6.2, we obtain $\mu(E) = \nu(E)$ for every Borel set E.

Corollary 6.4. Let $p = 1$ or $p = 2$. If μ is a Borel measure on $\mathbb{R}^N$, then

(a) $C_{\text{com}}(\mathbb{R}^N)$ is dense in $L^p(\mathbb{R}^N, \mu)$,

(b) the smallest closed subspace of $L^p(\mathbb{R}^N, \mu)$ containing all indicator functions of compact sets in $\mathbb{R}^N$ is $L^p(\mathbb{R}^N, \mu)$ itself.

REMARK. The scalars are assumed to be the same for $C_{\text{com}}(\mathbb{R}^N)$ as for $L^1(\mathbb{R}^N, \mu)$ and $L^2(\mathbb{R}^N, \mu)$; the corollary is valid both for real scalars and for complex scalars.

PROOF. If E is a Borel set of finite μ measure and if ϵ is given, Theorem 6.2 allows us to choose a compact set K with $K \subseteq E$ and $\mu(E - K) < \epsilon$. Then $\int_{\mathbb{R}^N} |I_E - I_K|^p\, d\mu = \mu(E - K) < \epsilon$, and consequently the closure in $L^p(\mathbb{R}^N)$ of the set of all indicator functions of compact sets contains all indicator functions of Borel sets of finite μ measure. Proposition 5.56 shows consequently that the smallest closed subspace of $L^p(\mathbb{R}^N)$ containing all indicator functions of compact sets is $L^p(\mathbb{R}^N)$ itself. This proves (b).

For (a), let K be compact, and use (iii) to choose a decreasing sequence $\{f_n\}$ of real-valued members of $C_{\text{com}}(\mathbb{R}^N)$ with pointwise limit I_K. Since f_1^p is integrable, dominated convergence yields $\lim_n \int_{\mathbb{R}^N} |f_n - I_K|^p\, d\mu = 0$. Hence the closure of $C_{\text{com}}(\mathbb{R}^N)$ in $L^p(\mathbb{R}^N)$ contains all indicator functions of compact sets. By Proposition 5.55d this closure contains the smallest closed subspace of $L^p(\mathbb{R}^N)$ containing all indicator functions of compact sets. Conclusion (b) shows that the latter subspace is $L^p(\mathbb{R}^N)$ itself. This proves (a).

Fix an integer $n \geq 0$, and let $(a_1, \dots, a_N)$ be an n-tuple of integers. The **diadic cube** $Q_n(a_1, \dots, a_N)$ in $\mathbb{R}^N$ of side 2^{-n} is defined to be the geometric rectangle

$$Q_n(a_1, \dots, a_N) = \big\{(x_1, \dots, x_N) \,\big|\, 2^{-n}a_j < x_j \leq 2^{-n}(a_j + 1) \text{ for } 1 \leq j \leq N\big\}.$$

Let $\mathcal{Q}_n$ be the set of all diadic cubes of side 2^{-n}. The members of $\mathcal{Q}_n$ are disjoint and have union $\mathbb{R}^N$. Thus we can associate uniquely to each x in $\mathbb{R}^N$ a sequence $\{Q_n\}$ of diadic cubes such that x is in Q_n and Q_n is in $\mathcal{Q}_n$. Since for each n, the members of $\mathcal{Q}_{n+1}$ are obtained by subdividing each member of $\mathcal{Q}_n$ into 2^N disjoint smaller diadic cubes, the diadic cubes Q_n associated to x must have the property that $Q_n \supseteq Q_{n+1}$ for all $n \geq 0$.

Lemma 6.5. Any open set in $\mathbb{R}^N$ is the countable disjoint union of diadic cubes.

PROOF. Let an open set U be given. We may assume that $U \neq \mathbb{R}^N$, so that $U^c \neq \varnothing$. We describe which diadic cubes to include in a collection $\mathcal{A}$ so that $\mathcal{A}$ has the required properties. If x is in U, then $D(x, U^c) = d$ is positive since U^c is closed and nonempty. Let $\{Q_n\}$ be the sequence of diadic cubes associated to x. The distance between any two points of Q_n is $\leq 2^{-n}\sqrt{N}$, and this is $< d$ if n is sufficiently large. Hence Q_n is contained in U for n sufficiently large. The cube in $\mathcal{A}$ that contains x is to be the Q_n with n as small as possible so that $Q_n \subseteq U$.

The construction has been arranged so that the union of the diadic cubes in $\mathcal{A}$ is exactly U. Suppose that Q and Q' are members of $\mathcal{A}$ obtained from respective points x and x' in U. If $Q \cap Q' \neq \varnothing$, let x'' be in the intersection. Then Q and Q' are two of the diadic cubes in the sequence associated to x'', and one has to contain the other. Without loss of generality, suppose that $Q \supseteq Q'$. Then x' lies in Q as well as Q', and we should have selected Q for x' rather than Q' if $Q \neq Q'$. We conclude that $Q = Q'$, and thus the members of $\mathcal{A}$ are disjoint. Each collection $\mathcal{Q}_n$ is countable, and therefore the collection $\mathcal{A}$ is countable.

Proposition 6.6. Any Borel measure on $\mathbb{R}^N$ is determined by its values on all the diadic cubes.

REMARK. We shall apply this result in the present section in connection with Lebesgue measure on $\mathbb{R}^N$ and in Section 8 in connection with general Borel measures on $\mathbb{R}^1$.

PROOF. The values on the diadic cubes determine the values on all open sets by Lemma 6.5, and the values on all open sets determine the values on all Borel sets by Theorem 6.2.

Corollary 6.7. There exists a unique Borel measure on $\mathbb{R}^N$ for which the measure of each geometric rectangle is the product of the lengths of the sides. The measure is the N-fold product of 1-dimensional Lebesgue measure.

REMARKS. The uniqueness is immediate from Proposition 6.6. The first version of Lebesgue measure that we constructed has the property stated in the corollary and therefore proves existence. All the other versions of **Lebesgue measure** we constructed have the same property, and so all such versions are equal. The corollary therefore allows us to use Fubini's Theorem for any decomposition $\mathbb{R}^N = \mathbb{R}^m \times \mathbb{R}^n$ with $m + n = N$. As in the 1-dimensional case, we shall often write dx for Lebesgue measure.

Corollary 6.7 gives one characterization of Lebesgue measure. We shall use Proposition 6.6 to give a second characterization, which will be in terms of translation invariance.

Proposition 6.8. Under a Borel function $F : \mathbb{R}^N \to \mathbb{R}^{N'}$, $F^{-1}(E)$ is in $\mathcal{B}_N$ whenever E is in $\mathcal{B}_{N'}$. In particular, this conclusion is valid if F is continuous.

PROOF. The set of E's for which $F^{-1}(E)$ is in $\mathcal{B}_N$ is a σ-algebra, and the result will follow if this set of E's contains the open geometric rectangles of $\mathbb{R}^{N'}$. If F_j denotes the j^{th} component of F, then $F_j : \mathbb{R}^N \to \mathbb{R}^1$ is Borel measurable and Proposition 5.6c shows that ${F_j}^{-1}(U_j)$ is a Borel set in $\mathbb{R}^N$ if U_j is open in $\mathbb{R}^1$. Then $F^{-1}(U_1 \times \cdots \times U_{N'}) = \bigcap_{j=1}^{N'} {F_j}^{-1}(U_j)$ is a Borel set in $\mathbb{R}^N$.

Corollary 6.9. Any homeomorphism of $\mathbb{R}^N$ carries $\mathcal{B}_N$ to $\mathcal{B}_N$.

Corollary 6.9 is a special case of Proposition 6.8. The particular homeomorphisms of interest at the moment are **translations** and **dilations**. Translation by x_0 is the homeomorphism $\tau_{x_0}(x) = x + x_0$. Its operation on a set E is given by $\tau_{x_0}(E) = \{\tau_{x_0}(x) \mid x \in E\} = \{x + x_0 \mid x \in E\} = E + x_0$, and its operation on a function f on $\mathbb{R}^N$ is given by $\tau_{x_0}(f)(x) = f(\tau_{x_0}^{-1}(x)) = f(x - x_0)$. Its operation on an indicator function I_E is $\tau_{x_0}(I_E)(x) = I_E(x - x_0) = I_{E+x_0}(x) = I_{\tau_{x_0}(E)}(x)$. Because of Corollary 6.9, translations operate on measures, the formula being $\tau_{x_0}(\mu)(E) = \mu(\tau_{x_0}^{-1}(E))$; since homeomorphisms carry compact sets to compact sets, the right side is a Borel measure if μ is a Borel measure. The actions of τ_{x_0} on functions and measures are related by integration. If $f \geq 0$ is a Borel function, then so is $\tau_{x_0}(f)$, and $\int_{\mathbb{R}^N} f\, d(\tau_{x_0}\mu) = \int_{\mathbb{R}^N} \tau_{x_0}^{-1}(f)\, d\mu$; this formula is verified by checking it for indicator functions and then passing to simple functions ≥ 0 by linearity and to Borel functions $f \geq 0$ by monotone convergence.

Dilation δ_c by a nonzero real c is given on members of $\mathbb{R}^N$ by $\delta_c(x) = cx$, and the operations on sets, functions, indicator functions, and measures are analogous to the corresponding operations for translations. Although dilations will play a recurring role in this book, the notation δ_c will be used only in the present section.

Theorem 6.10. Lebesgue measure m on $\mathbb{R}^N$ is translation invariant in the sense that $\tau_{x_0}(m) = m$ for every x_0 in $\mathbb{R}^N$. In fact, Lebesgue measure is the unique translation-invariant Borel measure on $\mathbb{R}^N$ that assigns measure 1 to the diadic cube $Q_0(0,\dots,0)$. The effect of dilations on Lebesgue measure is that $\delta_c(m) = |c|^{-N}m$, i.e., $\int_{\mathbb{R}^N} f(cx)\,dx = |c|^{-N}\int_{\mathbb{R}^N} f(x)\,dx$ for every nonnegative Borel function f.

REMARKS. From one point of view, translation and dilation are examples of bounded linear operators on each $L^p(\mathbb{R}^N, dx)$, with translation preserving norms and with dilation multiplying norms by a constant depending on p and the particular dilation. From another point of view, translation and dilation are especially simple examples of changes of variables. Operationally the theorem allows us to write $dy = dx$ when $y = \tau_{x_0}(x)$ and $dz = |c|^N\,dx$ when $z = cx$. These effects of translations and dilations on integration with respect to Lebesgue measure are special cases of the general change-of-variables formula to be proved in Section 5.

PROOF. For any x_0 in $\mathbb{R}^N$, m and $\tau_{x_0}(m)$ assign the product of the lengths of the sides as measure to any diadic cube. From Proposition 6.6 we conclude that $m = \tau_{x_0}(m)$. The assertion about the effect of dilations on Lebesgue measure is proved similarly.

We still have to prove the uniqueness. Let μ be a translation-invariant Borel measure. The members of $\mathcal{Q}_n$ are translates of one another and hence have equal μ measure. The members of $\mathcal{Q}_{n+1}$ are obtained by partitioning each member of $\mathcal{Q}_n$ into 2^N members of $\mathcal{Q}_{n+1}$ that are translates of one another. Thus the μ measure of any member of $\mathcal{Q}_{n+1}$ is 2^{-N} times the μ measure of any member of $\mathcal{Q}_n$. Consequently the μ measure of any diadic cube is completely determined by the value of μ on $Q_0(0,\dots,0)$, which is a member of $\mathcal{Q}_0$. The uniqueness then follows by another application of Proposition 6.6.

For a continuous function on a closed bounded interval, it was shown at the end of Section V.3 that the Riemann integral equals the Lebesgue integral. The next proposition gives an N-dimensional analog. A general comparison of the Riemann and Lebesgue integrals will be given in Section 4.

Proposition 6.11. For a continuous function on a compact geometric rectangle, the Riemann integral equals the Lebesgue integral.

PROOF. The two are equal in the 1-dimensional case, and the N-dimensional cases of each may be computed by iterated 1-dimensional integrals—as a result of

Corollary 3.33 in the case of the Riemann integral and as a result of the definition of Lebesgue measure as a product and the use of Fubini's Theorem (Theorem 5.47) in the case of the Lebesgue integral.

So far, we have worked in this section only with Lebesgue measure on the Borel sets. The **Lebesgue measurable sets** are those sets that occur when Lebesgue measure is completed. The Lebesgue measurable sets of measure 0 are of particular interest. In Section III.8 we defined an ostensibly different notion of measure 0 by saying that a set in $\mathbb{R}^N$ is of measure 0 if for any $\epsilon > 0$, it can be covered by a countable set of open geometric rectangles of total volume less than ϵ, and Theorem 3.29 characterized the Riemann integrable functions on a compact geometric rectangle as those functions whose discontinuities form a set of measure 0 in this sense. Later, Proposition 5.39 showed for $\mathbb{R}^1$ that a set has measure 0 in this sense if and only if it is Lebesgue measurable of Lebesgue measure 0. This equivalence extends to $\mathbb{R}^N$, as the next proposition shows.

Proposition 6.12. In $\mathbb{R}^N$, the Lebesgue measurable sets of measure 0 are exactly the subsets E of $\mathbb{R}^N$ with the following property: for any $\epsilon > 0$, the set E can be covered by countably many geometric rectangles of total volume less than ϵ.

PROOF. Let m be Lebesgue measure on $\mathbb{R}^N$. If E has the stated property, let E_n be the union of the given countable collection of geometric rectangles of total volume $< 1/n$ used to cover E. Proposition 5.1g shows that $m(E_n) < 1/n$, and hence the Borel set $E' = \bigcap_k E_k$ has $m(E') < 1/n$ for every n. Therefore $m(E') = 0$. Since $E \subseteq E'$, E is Lebesgue measurable and has Lebesgue measure 0.

Conversely if E is Lebesgue measurable of Lebesgue measure 0 and if $\epsilon > 0$ is given, we are to find a union of open geometric rectangles containing E and having total volume $< \epsilon$. Find a set E' in $\mathcal{B}_N$ with $E \subseteq E'$ and $m(E') = 0$. It is enough to handle E'. Writing $\mathbb{R}^N$ as the union of compact geometric cubes C_n of side $2n$ centered at the origin and covering $E' \cap C_n$ up to $\epsilon/2^n$, we see that we may assume that E' is bounded, being contained in some cube C_n.

Within $\mathbb{R}^1 \cap [-n, n]$, we know that the set of finite unions of intervals is an algebra $\mathcal{A}_1^{(n)}$ of sets such that $\mathcal{B}_1^{(n)} = \mathcal{B}_1 \cap [-n, n]$ is the smallest σ-algebra containing $\mathcal{A}_1^{(n)}$. Applying Proposition 5.40 inductively, we see that the set of finite disjoint unions of N-fold products of members of $\mathcal{A}_1^{(n)}$ is an algebra $\mathcal{A}_N^{(n)}$, and then Proposition 6.1 shows that the smallest σ-algebra containing $\mathcal{A}_N^{(n)}$ is $\mathcal{B}_N^{(n)} = \mathcal{B}_N \cap C_n$. Proposition 5.38 shows that the measure m on $\mathcal{B}_N^{(n)}$ is given by m^*, where $m^*(A)$ is the infimum of countable unions of members of $\mathcal{A}_N^{(n)}$ that cover A. Consequently the subset E' of C_n can be covered by countably many geometric rectangles of total volume $< \epsilon$. Doubling these rectangles about their centers and discarding their edges, we obtain a covering of E' by open rectangles of total volume $< 2^N \epsilon$, and we have the required covering.

Borel measurable sets have two distinct advantages over Lebesgue measurable sets. One advantage is that Borel measurable sets are independent of the particular Borel measure in question, whereas the sets in the completion of a σ-algebra relative to a Borel measure very much depend on the particular measure. The other advantage is that Fubini's Theorem applies in a tidy fashion to Borel measurable functions as a consequence of the identity $\mathcal{B}_m \times \mathcal{B}_n = \mathcal{B}_{m+n}$ given in Proposition 6.1. By contrast, there are Lebesgue measurable sets for $\mathbb{R}^N$ that are not in the product of the σ-algebras of Lebesgue measurable sets from $\mathbb{R}^m$ and $\mathbb{R}^{N-m}$. For example, take a set E in $\mathbb{R}^1$ that is not Lebesgue measurable; such a set is produced in Problem 1 at the end of the present chapter. Then $E \times \{0\}$ in $\mathbb{R}^2$ is a subset of the Borel set $\mathbb{R}^1 \times \{0\}$, and hence it is Lebesgue measurable of measure 0. However, $E \times \{0\}$ is not in the product σ-algebra, because a section of a function measurable with respect to the product has to be measurable with respect to the appropriate factor (Lemma 5.46).

On the other hand, Lebesgue measurable functions are sometimes unavoidable. An example occurs with Riemann integrability: In view of Proposition 6.12, Theorem 3.29 says that the Riemann integrable functions on a compact geometric rectangle are exactly the functions whose discontinuities form a Lebesgue measurable set of Lebesgue measure 0, and Problems 31–33 at the end of Chapter V produced such a function in the 1-dimensional case that is not a Borel measurable function.

The upshot is that a little care is needed when using Fubini's Theorem and Lebesgue measurable sets at the same time, and there are times when one wants to do so. The situation is a little messy but not intractable. Problem 12 at the end of Chapter V showed that a Lebesgue measurable function can be adjusted on a set of Lebesgue measure 0 so as to become Borel measurable. Using this fact, one can write down a form of Fubini's Theorem for Lebesgue measurable functions that is usable even if inelegant.

2. Convolution

Convolution is an important operation available for functions on $\mathbb{R}^N$. On a formal level, the **convolution** $f * g$ of two functions f and g is

$$(f * g)(x) = \int_{\mathbb{R}^N} f(x-y)g(y)\,dy.$$

One place convolution arises is as a limit of a linear combination of translates: We shall see in Proposition 6.13 that the convolution at x may be written also as $\int_{\mathbb{R}^N} f(y)g(x-y)\,dy$. If f is fixed and if finite sets of translation operators τ_{y_i} and of weights $f(y_i)$ are given, then the value at x of the linear combination $\sum_i f(y_i)\tau_{y_i}$ applied to g and evaluated at x is $\sum_i f(y_i)g(x-y_i)$. Corollary 6.17 will show a sense in which we can think of $\int_{\mathbb{R}^N} f(y)g(x-y)\,dy$ as a limit of such expressions.

To make mathematical sense out of $f*g$, let us begin with the case that f and g are nonnegative Borel functions on $\mathbb{R}^N$. The assertion is that $f*g$ is meaningful as a Borel function ≥ 0. In fact, $(x,y)\mapsto f(x-y)$ is the composition of the continuous function $F:\mathbb{R}^{2N}\to\mathbb{R}^N$ given by $F(x,y)=x-y$, followed by the Borel function $f:\mathbb{R}^N\to[0,+\infty]$. If U is open in $[0,+\infty]$, then $f^{-1}(U)$ is in $\mathcal{B}_N$, and Proposition 6.8 shows that $(f\circ F)^{-1}(U)=F^{-1}(f^{-1}(U))$ is in $\mathcal{B}_{2N}$. Then the product $(x,y)\mapsto f(x-y)g(y)$ is a Borel function, and Fubini's Theorem (Theorem 5.47) and Proposition 6.1 combine to show that $x\mapsto(f*g)(x)$ is a Borel function ≥ 0.

Proposition 6.13. For nonnegative Borel functions on $\mathbb{R}^N$,

(a) $f*g=g*f$,
(b) $f*(g*h)=(f*g)*h$.

PROOF. We use Theorem 6.10 for both parts and also Fubini's Theorem for (b). For (a), the changes of variables $y\mapsto y+x$ and then $y\mapsto -y$ give $\int_{\mathbb{R}^N} f(x-y)g(y)\,dy=\int_{\mathbb{R}^N} f(-y)g(y+x)\,dx=\int_{\mathbb{R}^N} f(y)g(x-y)\,dy$. For (b), the computation is

$$\begin{aligned}(f*(g*h))(x)&=\int_{\mathbb{R}^N} f(x-y)(g*h)(y)\,dy\\&=\int_{\mathbb{R}^N}\Big[\int_{\mathbb{R}^N} f(x-y)g(y-z)h(z)\,dz\Big]\,dy\\&=\int_{\mathbb{R}^N}\Big[\int_{\mathbb{R}^N} f(x-y)g(y-z)h(z)\,dy\Big]\,dz\\&=\int_{\mathbb{R}^N}\Big[\int_{\mathbb{R}^N} f(x-z-y)g(y)h(z)\,dy\Big]\,dz\\&=\int_{\mathbb{R}^N}(f*g)(x-z)h(z)\,dz=((f*g)*h)(x),\end{aligned}$$

the change of variables $y\mapsto y+z$ being used for the fourth equality.

In order to have a well-defined expression for $f*g$ when f and g are not necessarily ≥ 0, we need conditions under which the nonnegative case leads to something finite. The conditions we use ensure finiteness of $(|f|*|g|)(x)$ for almost every x. For real-valued f and g, we then define $f*g(x)$ by subtraction at the points where $(|f|*|g|)(x)$ is finite, and we define it to be 0 elsewhere. For complex-valued f and g, we define $(f*g)(x)$ as a linear combination of the appropriate parts where $(|f| * |g|)(x)$ is finite, and we define it to be 0 elsewhere. When we proceed this way, the commutativity and associativity properties in Proposition 6.13 will be valid even though f and g are not necessarily ≥ 0.

Proposition 6.14. For nonnegative Borel functions f and g on $\mathbb{R}^N$, convolution is finite almost everywhere in the following cases, and then the indicated inequalities of norms are satisfied:

(a) for f in $L^1(\mathbb{R}^N)$ and g in $L^1(\mathbb{R}^N)$, and then $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$,
(b) for f in $L^1(\mathbb{R}^N)$ and g in $L^2(\mathbb{R}^N)$, and then $\|f * g\|_2 \leq \|f\|_1 \|g\|_2$,
for f in $L^2(\mathbb{R}^N)$ and g in $L^1(\mathbb{R}^N)$, and then $\|f * g\|_2 \leq \|f\|_2 \|g\|_1$,
(c) for f in $L^1(\mathbb{R}^N)$ and g in $L^\infty(\mathbb{R}^N)$, and then $\|f * g\|_\infty \leq \|f\|_1 \|g\|_\infty$,
for f in $L^\infty(\mathbb{R}^N)$ and g in $L^1(\mathbb{R}^N)$, and then $\|f * g\|_\infty \leq \|f\|_\infty \|g\|_1$,
(d) for f in $L^2(\mathbb{R}^N)$ and g in $L^2(\mathbb{R}^N)$, and then $\|f * g\|_\infty \leq \|f\|_2 \|g\|_2$.

Consequently $f * g$ is defined in the above situations even if the scalar-valued functions f and g are not necessarily ≥ 0, and the estimates on the norm of $f * g$ are still valid.

PROOF. For (a) and the first conclusions in (b) and (c), let p be $1, 2$, or ∞ as appropriate. By Minkowski's inequality for integrals (Theorem 5.60),

$$\begin{aligned}\|f * g\|_p &= \Big\| \int_{\mathbb{R}^N} f(y) g(x-y)\, dy \Big\|_{p,x} \leq \int_{\mathbb{R}^N} \|f(y) g(x-y)\|_{p,x}\, dy \\ &= \int_{\mathbb{R}^N} |f(y)|\, \|g(x-y)\|_{p,x}\, dy = \int_{\mathbb{R}^N} |f(y)|\, \|g\|_p\, dy = \|f\|_1 \|g\|_p,\end{aligned}$$

the next-to-last equality following from the translation invariance of dx. The second conclusions in (b) and (c) require only notational changes.

For (d), we have

$$\begin{aligned}\sup_x |(f * g)(x)| &= \sup_x \Big| \int_{\mathbb{R}^N} f(y) g(x-y)\, dy \Big| \\ &\leq \sup_x \|f\|_2 \|g(x-y)\|_{2,y} = \|f\|_2 \|g\|_2,\end{aligned}$$

the inequality following from the Schwarz inequality and the last step following from translation invariance of dy and invariance under $y \mapsto -y$.

Going over these arguments, we see that we may use them even if f and g are not necessarily ≥ 0. Then the last statement of the proposition follows.

Next let us relate the translation operators of Section 1 to convolution. The formula for the effect of a translation operator on a function is $\tau_t(f)(x) = f(x-t)$.

Proposition 6.15. Convolution commutes with translations in the sense that $\tau_t(f * g) = (\tau_t f) * g = f * \tau_t g$.

PROOF. It is enough to treat functions ≥ 0. Then we have $\tau_t(f * g)(x) = (f * g)(x-t) = \int_{\mathbb{R}^N} f(x-t-y) g(y)\, dy$, which equals $\int_{\mathbb{R}^N} (\tau_t f)(x-y) g(y)\, dy = ((\tau_t f) * g)(x)$ on the one hand and, because of translation invariance of Lebesgue measure, equals $\int_{\mathbb{R}^N} f(x-y) g(y-t)\, dy = (f * \tau_t g)(x)$ on the other hand.

Proposition 6.16. If $p = 1$ or $p = 2$, then translation of a function is continuous in the translation parameter in $L^p(\mathbb{R}^N, dx)$. In other words, if f is in L^p relative to Lebesgue measure, then $\lim_{h\to 0} \|\tau_{t+h} f - \tau_t f\|_p = 0$ for all t.

REMARK. However, continuity fails on L^∞. In this case, there is a substitute result, and we take that up in a moment.

PROOF. Let f be in L^p. By translation invariance of Lebesgue measure, $\|\tau_{t+h} f - \tau_t f\|_p = \|\tau_h f - f\|_p$. If g is in $C_{\text{com}}(\mathbb{R}^N)$, then $\|\tau_h g - g\|_p^p = \int_{\mathbb{R}^N} |g(x-h) - g(x)|^p\, dx$, and dominated convergence shows that this tends to 0 as h tends to 0. Let $\epsilon > 0$ and f be given. By Corollary 6.4a, $C_{\text{com}}(\mathbb{R}^N)$ is dense in $L^p(\mathbb{R}^N, dx)$, and thus we can choose g in $C_{\text{com}}(\mathbb{R}^N)$ with $\|f - g\|_p < \epsilon$. Then

$$\begin{aligned}\|\tau_h f - f\|_p &\le \|\tau_h f - \tau_h g\|_p + \|\tau_h g - g\|_p + \|g - f\|_p \\ &= 2\|f - g\|_p + \|\tau_h g - g\|_p \le 2\epsilon + \|\tau_h g - g\|_p.\end{aligned}$$

If h is close enough to 0, the term $\|\tau_h g - g\|_p$ is $< \epsilon$, and then $\|\tau_h f - f\|_p < 3\epsilon$.

Corollary 6.17. Let $p = 1$ or $p = 2$, and let $g_1, \ldots, g_r$ be finitely many functions in $L^p(\mathbb{R}^N)$. If a positive number ϵ and a function f in $L^1(\mathbb{R}^N)$ are given, then there exist finitely many members y_j of $\mathbb{R}^N$, $1 \le j \le n$, and constants c_j such that $\big\| f * g_k - \sum_{j=1}^n c_j \tau_{y_j} g_k \big\|_p < \epsilon$ for $1 \le k \le r$.

REMARK. In the case $r = 1$, the corollary says that any convolution $f * g$ can be approximated in L^p by a linear combination of translates of g. The result will be used in Chapter VIII with $r > 1$.

PROOF. Let V be the set of functions f in $L^1(\mathbb{R}^N)$ for which this kind of approximation is possible for every $\epsilon > 0$. The main step is to show that V contains the indicator functions of the compact sets in $\mathbb{R}^N$. Let K be compact, and let I_K be its indicator function. Proposition 6.16 shows that the functions $y \mapsto \tau_y g_k$ are continuous from K into $L^p(\mathbb{R}^N)$ for $1 \le k \le r$, and therefore these functions are uniformly continuous. Fix $\epsilon > 0$, and let $\delta > 0$ be such that $\|\tau_y g_k - \tau_{y'} g_k\|_p < \epsilon$ for all k whenever $|y - y'| < \delta$ and y and y' are in K. For each y in K, form the open ball $B(\delta; y)$ in $\mathbb{R}^N$. These balls cover K, and finitely many suffice; let their centers be $y_1, \ldots, y_n$. Define sets $S_1, \ldots, S_n$ inductively as follows: S_j is the subset of K where $|y - y_j| < \delta$ but $|y - y_i| \ge \delta$ for $i < j$. Then $K = \bigcup_{j=1}^n S_j$ disjointly. By the choice of δ, we have $\|\tau_y g_k - \tau_{y_j} g_k\|_p < \epsilon$ for all y in S_j and all k. Using Minkowski's inequality for integrals (Theorem 5.60), and writing m for Lebesgue measure, we have

$$\begin{aligned}\big\| I_{S_j} * g_k - m(S_j)\tau_{y_j} g_k \big\|_p &= \big\| \textstyle\int_{S_j} (g_k(x-y) - g_k(x-y_j))\, dy \big\|_p \\ &\le \textstyle\int_{S_j} \|g_k(x-y) - g_k(x-y_j)\|_{p,x}\, dy \\ &\le \epsilon m(S_j).\end{aligned}$$

Summing over j gives

$$\left\| I_K * g_k - \sum_{j=1}^n m(S_j)\tau_{y_j} g_k \right\|_p \le \epsilon m(K).$$

Since ϵ is arbitrary, I_K lies in V.

If f_1 and f_2 are in V and if $g_1, \dots, g_r$ are given, then we may assume, by taking the union of the sets of members y_j of $\mathbb{R}^N$ and by setting any unnecessary constants c_j equal to 0, that the translates used for f_1 and f_2 with the same $\epsilon > 0$ are the same. Thus we can write $\left\| f_1 * g_k - \sum_{j=1}^n c_j \tau_{y_j} g_k \right\|_p < \epsilon/2$ and $\left\| f_2 * g_k - \sum_{j=1}^n d_j \tau_{y_j} g_k \right\|_p < \epsilon/2$ for suitable y_j's and c_j's, and the triangle inequality gives $\left\| (f_1 + f_2) * g_k - \sum_{j=1}^n (c_j + d_j)\tau_{y_j} g_k \right\|_p < \epsilon$. Hence V is closed under addition. Similarly V is closed under scalar multiplication. If $f_l \to f$ in L^1 with f_l in V and if $\epsilon > 0$ is given, choose l large enough so that $\|f - f_l\|_1 < \epsilon \big/ (2 \max \|g_k\|_p)$. If $\left\| f_l * g_k - \sum_{j=1}^n c_j^{(l)} \tau_{y_j^{(l)}} g_k \right\|_p < \epsilon/2$, then the inequality $\|f * g_k - f_l * g_k\|_p \le \|f - f_l\|_1 \|g_k\|_p$ and the triangle inequality together give $\left\| f * g_k - \sum_{j=1}^n c_j^{(l)} \tau_{y_j^{(l)}} g_k \right\|_p < \epsilon$. Hence f is in V, and V is closed. By Corollary 6.4b, $V = L^1(\mathbb{R}^N)$, and the proof is complete.

In some cases with $L^\infty(\mathbb{R}^N)$, results have more content when phrased in terms of the supremum norm $\|f\|_{\sup} = \sup_{x \in \mathbb{R}^N} |f(x)|$ defined in Section V.9. For a continuous function f, the two norms agree because the set where $|f(x)| > M$ is open and therefore has positive measure if it is nonempty. For a bounded function f, the condition $\lim_{h \to 0} \|\tau_h f - f\|_{\sup} = 0$ is equivalent to uniform continuity of f, basically by definition. The functions f in L^∞ for which $\lim_{h \to 0} \|\tau_h f - f\|_\infty = 0$ are not much more general than the bounded uniformly continuous functions; we shall see shortly that they can be adjusted on a set of measure 0 so as to be bounded and uniformly continuous.

Proposition 6.18. In $\mathbb{R}^N$ with Lebesgue measure, the convolution of L^1 with L^∞, or of L^∞ with L^1, or of L^2 with L^2 results in an everywhere-defined bounded uniformly continuous function, not just an L^∞ function. Moreover,

$$\|f * g\|_{\sup} \le \|f\|_1 \|g\|_\infty, \quad \|f * g\|_{\sup} \le \|f\|_\infty \|g\|_1, \quad \text{or} \quad \|f * g\|_{\sup} \le \|f\|_2 \|g\|_2$$

in the various cases.

PROOF. We give the proof when f is in L^1 and g is in L^∞, the other cases being handled similarly. The bound follows from the computation $\|f * g\|_{\sup} = \sup_x \left| \int_{\mathbb{R}^N} f(x - y) g(y)\, dy \right| \le \sup_x \|g\|_\infty \int_{\mathbb{R}^N} |f(x - y)|\, dy = \|f\|_1 \|g\|_\infty$.

For uniform continuity we use Proposition 6.15 and the bound $\|f * g\|_{\sup} \le \|f\|_1 \|g\|_\infty$ to make the estimate

$$\begin{aligned}\|\tau_h(f * g) - (f * g)\|_{\sup} &= \|(\tau_h f) * g - f * g\|_{\sup} \\ &= \|(\tau_h f - f) * g\|_{\sup} \le \|\tau_h f - f\|_1 \|g\|_\infty,\end{aligned}$$

and then we apply Proposition 6.16 to see that the right side tends to 0 as h tends to 0.

A corollary of Proposition 6.18 gives a first look at how differentiability interacts with convolution.

Corollary 6.19. Suppose that f is a compactly supported function of class C^n on $\mathbb{R}^N$ and that g is in $L^p(\mathbb{R}^N, dx)$ with p equal to $1, 2,$ or ∞. Then $f * g$ is of class C^n, and $D(f * g) = (Df) * g$ for any iterated partial derivative of order $\le n$.

PROOF. First suppose that $n = 1$. Fix j with $1 \le j \le N$, and put $D_j = \partial/\partial x_j$. The function $(D_j f) * g$ is continuous by Proposition 6.18. If we can prove that $D_j(f * g)(x)$ exists and equals $((D_j f) * g)(x)$ for each x, then it will follow that $D_j(f * g)$ is continuous. This fact for all j implies that $f * g$ is of class C^1, by Theorem 3.7, and the result for $n = 1$ will have been proved. The result for higher n can then be obtained by iterating the result for $n = 1$.

Thus we are to prove that $D_j(f * g)(x)$ exists and equals $((D_j f) * g)(x)$ for each x. In the respective cases $p = 1, 2, \infty$, put $p' = \infty, 2, 1$. Let e_j be the j^{th} standard basis vector of $\mathbb{R}^N$ and let h be real with $|h| \le 1$. Proposition 6.15 gives

$$h^{-1}\big((f * g)(x + he_j) - (f * g)(x)\big) = \big((h^{-1}(\tau_{-he_j} f - f)) * g\big)(x). \qquad (*)$$

Proposition 3.28a shows that $h^{-1}(\tau_{-he_j} f - f)$ converges uniformly, as $h \to 0$, to $D_j f$ on any compact set; since the support is compact, $h^{-1}(\tau_{-he_j} f - f)$ converges uniformly to $D_j f$ on $\mathbb{R}^N$. Hence the convergence occurs in L^∞, and dominated convergence shows that it occurs in L^1 and L^2 also. Combining Proposition 6.18 and $(*)$, we see that

$$|h^{-1}\big((f*g)(x+he_j)-(f*g)(x)\big)-(D_j f)(x)| \le \|h^{-1}(\tau_{-he_j} f - f) - D_j f\|_{p'} \|g\|_p.$$

The right side tends to 0 as $h \to 0$, and thus indeed $D_j(f * g)(x)$ exists and equals $((D_j f) * g)(x)$.

Twice in Chapter I we made use of an "approximate identity" in $\mathbb{R}^1$, a system of functions peaking at the origin such that convolution by these functions acts more and more like the identity operator on some class of functions. The first occasion of this kind was in Section I.9 in connection with the Weierstrass Approximation Theorem, where the functions in the system were $\varphi_n(x) = c_n(1-x^2)^n$ on $[-1, 1]$ with the constants c_n chosen to make the total integral be 1. The polynomials φ_n had the properties

(i) $\varphi_n(x) \geq 0$,
(ii) $\int_{-1}^{1} \varphi_n(x)\,dx = 1$,
(iii) for any $\delta > 0$, $\sup_{\delta \leq |x| \leq 1} \varphi_n(x)$ tends to 0 as n tends to infinity,

and the convolutions were with continuous functions f such that $f(0) = f(1) = 0$ and f vanishes outside $[0, 1]$. The second occasion was in Section I.10 in connection with Fejér's Theorem, where the functions in the system were trigonometric polynomials $K_N(x)$ such that

(i) $K_N(x) \geq 0$,
(ii) $\frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(x)\,dx = 1$,
(iii) for any $\delta > 0$, $\sup_{\delta \leq |x| \leq \pi} K_N(x)$ tends to 0 as n tends to infinity.

In this case the convolutions were with periodic functions of period 2π over an interval of length 2π, and the integrations involved $\frac{1}{2\pi}\,dx$ instead of dx.

Now we shall use the dilations of a single function in order to produce a more robust kind of approximate identity, this time on $\mathbb{R}^N$. One sense in which convolution by this system acts more and more like the identity appears in Theorem 6.20 below, and a sample application appears in Corollary 6.21. The corollary will illustrate how one can use an approximate identity to pass from conclusions about nice functions in some class to conclusions about all functions in the class.

Theorem 6.20. Let φ be in $L^1(\mathbb{R}^N, dx)$, not necessarily ≥ 0. Define

$$\varphi_\varepsilon(x) = \varepsilon^{-N}\varphi(\varepsilon^{-1}x) \qquad \text{for } \varepsilon > 0,$$

and put $c = \int_{\mathbb{R}^N} \varphi(x)\,dx$. Then the following hold:

(a) if $p = 1$ or $p = 2$ and if f is in $L^p(\mathbb{R}^N, dx)$, then

$$\lim_{\varepsilon \downarrow 0} \|\varphi_\varepsilon * f - cf\|_p = 0,$$

(b) the conclusion in (a) is valid for $p = \infty$ if f is in $L^\infty(\mathbb{R}^N, dx)$ and $\lim_{t \to 0} \|\tau_t f - f\|_\infty = 0$,
(c) if f is bounded on $\mathbb{R}^N$ and is continuous at x, then $\lim_{\varepsilon \downarrow 0}(\varphi_\varepsilon * f)(x) = cf(x)$,
(d) the convergence in (c) is uniform for any set E of x's such that f is uniformly continuous at the points of E.

Corollary 6.21. If f in $L^\infty(\mathbb{R}^N, dx)$ satisfies $\lim_{t\to 0} \|\tau_t f - f\|_\infty = 0$, then f can be adjusted on a set of measure 0 so as to be uniformly continuous.

PROOF. Let φ be a member of $C_{\text{com}}(\mathbb{R}^N)$ such that $\int_{\mathbb{R}^N} \varphi(x)\,dx = 1$. Fix a sequence $\{\varepsilon_n\}$ decreasing to 0 in $\mathbb{R}^1$. Proposition 6.18 shows that each $\varphi_{\varepsilon_n} * f$ is bounded and uniformly continuous for every n, and Theorem 6.20 shows that $\{\varphi_{\varepsilon_n} * f\}$ is Cauchy in L^∞. Since the L^∞ and supremum norms coincide for continuous functions, $\{\varphi_{\varepsilon_n} * f\}$ is uniformly Cauchy and must therefore be uniformly convergent. Let g be the limit function, which is necessarily bounded and uniformly continuous. Then $\|f-g\|_\infty \le \|f-\varphi_{\varepsilon_n} * f\|_\infty + \|\varphi_{\varepsilon_n} * f - g\|_\infty$, and both terms on the right tend to 0 as n tends to infinity. Consequently $\|f-g\|_\infty = 0$, and g is a bounded uniformly continuous function that differs from f only on a set of measure 0.

3. Borel Measures on Open Sets

A number of results in Sections 1–2 about Borel measures on $\mathbb{R}^N$ extend to suitably defined Borel measures on arbitrary nonempty open subsets V of $\mathbb{R}^N$, and we shall collect some of these results here in order to do two things: to prepare for the proof in Section 5 of the change-of-variables formula for the Lebesgue integral in $\mathbb{R}^N$ and to provide motivation for the treatment in Chapter XI of Borel measures on locally compact Hausdorff spaces.

Throughout this section, let V be a nonempty open subset of $\mathbb{R}^N$. We shall make use of the following lemma that generalizes to V three properties (i–iii) listed for $\mathbb{R}^N$ before Theorem 6.2 and Corollary 6.3. Let $C_{\text{com}}(V)$ be the vector space of scalar-valued continuous functions on V of compact support in V. If nothing is said to the contrary, the scalars may be either real or complex.

Lemma 6.22.

(a) There exists a sequence $\{F_n\}_{n=1}^\infty$ of compact subsets of V with union V such that $F_n \subseteq F_{n+1}^o$ for all n.

(b) For any compact subset K of V, there exists a decreasing sequence of open sets U_n with compact closure in V such that $\bigcap_{n=1}^\infty U_n = K$.

(c) For any compact subset K of V, there exists a decreasing sequence of functions in $C_{\text{com}}(V)$ with values in [0, 1] and with pointwise limit the indicator function of K.

PROOF. In (a), the case $V = \mathbb{R}^N$ was handled by (i) before Theorem 6.2. For $V \ne \mathbb{R}^N$, we can take $F_n = \big\{x \in V \,\big|\, D(x, V^c) \ge 1/n \text{ and } |x| \le n\big\}$ as long as n is $\ge$ some suitable n_0. We complete the definition of the F_n's by taking $F_1 = \cdots = F_{n_0-1} = F_{n_0}$.

In (b), the case $V = \mathbb{R}^N$ was handled by (ii) before Theorem 6.2. For $V \neq \mathbb{R}^N$, every x in K has $D(x, V^c) > 0$ since V^c is closed and is disjoint from K. The function $D(\,\cdot\,, V^c)$ is continuous and therefore has a positive minimum on K. Choose n_0 such that $D(x, V^c) \geq 1/n_0$ for x in K, i.e., $|x - y| \geq 1/n_0$ for all $x \in K$ and $y \in V^c$. Then $D(y, K) \geq 1/n_0$ if y is not in V. Let $U_n = \{y \in \mathbb{R}^N \mid D(y, K) < 1/n\}$ for $n > n_0$. This is an open set containing K, and its closure in $\mathbb{R}^N$ is contained in the set where $D(y, K) \leq 1/n$, which in turn is contained in V. The set where $D(y, K) \leq 1/n$ is closed and bounded in $\mathbb{R}^N$ and hence is compact. Therefore U_n^{cl} is contained in a compact subset of V. We complete the definition of the U_n's by letting $U_1, \ldots, U_{n_0}$ all equal U_{n_0+1}.

For (c), we argue as with (iii) before Corollary 6.3. Choose open sets U_n as in (b) that decrease and have intersection K, and apply Proposition 2.30e to obtain continuous functions $f_n : \mathbb{R}^N \to [0, 1]$ such that f_n is 1 on K and is 0 on U_n^c. The support of the function f_n is then contained in U_n^{cl}, which is compact. By replacing the functions f_n by $g_n = \min\{f_1, \ldots, f_n\}$, we may assume that they are pointwise decreasing. Then (c) follows.

The **Borel sets** in the open set V are the sets in the σ-algebra

$$\mathcal{B}_N(V) = \mathcal{B}_N \cap V = \{E \cap V \mid E \in \mathcal{B}_N\}$$

of subsets of V. We can regard V as a metric space by restricting the distance function on $\mathbb{R}^N$, and because V is open, the open sets of V are the open sets of $\mathbb{R}^N$ that are subsets of V. We shall prove the following proposition about these Borel sets after first proving a lemma.

Proposition 6.23. The σ-algebra $\mathcal{B}_N(V)$ is the smallest σ-algebra for V containing the open sets of V, and it is the smallest σ-algebra for V containing the compact sets of V.

Lemma 6.24. Let X be a nonempty set, let $\mathcal{U}$ be a family of subsets of X, let $\mathcal{B}$ be the smallest σ-ring of subsets of X containing $\mathcal{U}$, and let E be a member of $\mathcal{B}$. Then $\mathcal{B} \cap E$ is the smallest σ-ring containing $\mathcal{U} \cap E$.

PROOF OF LEMMA 6.24. Let $\mathcal{A}$ be the smallest σ-ring containing $\mathcal{U} \cap E$, and let $\mathcal{A}'$ be the smallest σ-ring containing $\mathcal{U} \cap E^c$. Since $\mathcal{B} \cap E$ is a σ-ring of subsets of X containing $\mathcal{U} \cap E$, $\mathcal{A}$ is contained in $\mathcal{B} \cap E$. Similarly $\mathcal{A}' \subseteq \mathcal{B} \cap E^c$. Thus the set of unions $A \cup A'$ with $A \in \mathcal{A}$ and $A' \in \mathcal{A}'$ is contained in $\mathcal{B}$, contains $\mathcal{U}$, and is closed under countable unions. To see that it is closed under differences, let $A_1 \cup A_1'$ and $A_2 \cup A_2'$ be such unions. Then $(A_1 \cup A_1') - (A_2 \cup A_2') = (A_1 - A_2) \cup (A_1' - A_2')$ exhibits the difference of the given sets as such a union. Hence the set of such unions is a σ-ring and must equal $\mathcal{B}$.

PROOF OF PROPOSITION 6.23. The statement about open sets follows from Lemma 6.24 by taking X to be $\mathbb{R}^N$, $\mathcal{U}$ to be the set of open sets in $\mathbb{R}^N$, and E to be V. The set $\mathcal{U} \cap E$ is the set of open subsets of V, and the lemma says that the smallest σ-ring containing $\mathcal{U} \cap E$ is $\mathcal{B}_N(V)$. This is a σ-algebra of subsets of V since V itself is in $\mathcal{U} \cap V$.

Let $\{F_n\}$ be the sequence of compact subsets of V produced by Lemma 6.22a. Since $V = \bigcup_{n=1}^{\infty} F_n$, V is a member of the smallest σ-ring of subsets containing the compact subsets of V. If F is a relatively closed subset of V, then each $F \cap F_n$ is compact, and the countable union F is therefore in this σ-algebra. Taking complements, we see that every open subset of V is in the smallest σ-algebra of subsets of V containing the compact sets. Therefore $\mathcal{B}_N(V)$ is contained in this σ-algebra and must equal this σ-algebra.

A **Borel function** on V is a scalar-valued function measurable with respect to $\mathcal{B}_N(V)$. A **Borel measure** on V is a measure on $\mathcal{B}_N(V)$ that is finite on every compact set in V.

Theorem 6.25. Every Borel measure μ on the nonempty open subset V of $\mathbb{R}^N$ is **regular** in the sense that the value of μ on any Borel set E in V is given by

$$\mu(E) = \sup_{\substack{K \subseteq E, \\ K \text{ compact in } V}} \mu(K) = \inf_{\substack{U \supseteq E, \\ U \text{ open in } V}} \mu(U).$$

REMARK. If μ is a Borel measure on V and if we define $\nu(E) = \mu(E \cap V)$ for Borel sets E of $\mathbb{R}^N$, then ν is a measure on the Borel sets of $\mathbb{R}^N$, but ν need not be finite on compact sets. Thus Theorem 6.25 is not a special case of Theorem 6.2.

PROOF. This is proved from parts (a) and (b) of Lemma 6.22 in exactly the same way that Theorem 6.2 is proved from items (i) and (ii) before the statement of that theorem.

Corollary 6.26. If μ and ν are Borel measures on V such that $\int_V f\,d\mu = \int_V f\,d\nu$ for all f in $C_{\text{com}}(V)$, then $\mu = \nu$.

PROOF. This is proved from Theorem 6.25 and Lemma 6.22c in the same way that Corollary 6.3 is proved from Theorem 6.2 and item (iii) before the statement of that corollary.

Corollary 6.27. Let $p = 1$ or $p = 2$. If μ is a Borel measure on V, then

(a) $C_{\text{com}}(V)$ is dense in $L^p(V, \mu)$,
(b) the smallest closed subspace of $L^p(V, \mu)$ containing all indicator functions of compact subsets of V is $L^p(V, \mu)$ itself,

(c) $C_{\text{com}}(V)$, as a normed linear space under the supremum norm, is separable,
(d) $L^p(V, \mu)$ is separable.

PROOF. Conclusions (a) and (b) are proved from Lemma 6.22c with the aid of Propositions 5.56 and 5.55d in the same way that Corollary 6.4 is proved from item (iii) before Corollary 6.3 with the aid of those propositions.

For (c), Lemma 6.22a produces a sequence $\{F_n\}_{n=1}^{\infty}$ of compact subsets of V with union V such that $F_n \subseteq F_{n+1}^o$ for all n. Since the open sets F_n^o cover V, they cover any compact subset K of V, and K must be contained in some F_n^o. Let us put that observation aside for a moment. For each n, we can identify the vector subspace of $C_{\text{com}}(V)$ of functions supported in F_n^o with a vector subspace of $C(F_n)$. The latter is separable by Corollary 2.59, and hence the vector subspace of $C_{\text{com}}(V)$ of functions supported in F_n^o is separable. If we form the union on n of these countable dense sets of certain vector subspaces and if we take into account that the functions supported in any compact subset of V have compact support within some F_n^o, we see that $C_{\text{com}}(V)$ is separable.

For (d), we apply (a) and (c) with V replaced by F_n^o and take into account that $\mu(F_n) < \infty$. Let f be arbitrary in $L^p(F_n^o, \mu\big|_{F_n^o})$. If $\epsilon > 0$ is given, choose g in $C_{\text{com}}(F_n^o)$ with $\|f - g\|_p \leq \epsilon$. Then choose h in the countable dense set D_n of $C_{\text{com}}(F_n^o)$ such that $\|g - h\|_{\sup} \leq \epsilon$. Since $\|f - h\|_p \leq \|f - g\|_p + \|g - h\|_p$ and $\|g - h\|_p^p = \int_{F_n^o} |g(x) = h(x)|^p \, d\mu(x) \leq \epsilon^p \mu(F_n)$, we obtain $\|f - h\|_p \leq \epsilon + \epsilon\mu(F_n)^{1/p}$. Hence the closure in $L^p(V, \mu)$ of the countable set $D = \bigcup_{n=1}^{\infty} D_n$ contains f. In particular, D^{cl} is a vector subspace containing all indicator functions of compact subsets of V. By (b), $D^{\text{cl}} = L^p(V, \mu)$.

Proposition 6.28. With V still open in $\mathbb{R}^N$, let V' be an open set in $\mathbb{R}^{N'}$. Under a continuous function or even a Borel measurable function $F : V \to V'$, $F^{-1}(E)$ is in $\mathcal{B}_N(V)$ whenever E is in $\mathcal{B}_{N'}(V')$.

PROOF. The set of E's for which $F^{-1}(E)$ is in $\mathcal{B}_N(V)$ is a σ-algebra, and this σ-algebra contains the open geometric rectangles of $\mathbb{R}^{N'}$ by the same argument as for Proposition 6.8. Thus it contains $\mathcal{B}_{N'}(V')$.

Corollary 6.29. If V' is a second nonempty open subset in $\mathbb{R}^N$ besides V, then any homeomorphism of V onto V' carries $\mathcal{B}_N(V)$ to $\mathcal{B}_N(V')$.

If K is a nonempty compact subset of $\mathbb{R}^N$, it will be convenient to be able to speak of the Borel sets in K, just as we can speak of the Borel sets in an open subset V of $\mathbb{R}^N$. The theory for K is easier than the theory for V, partly because Borel measures on K can all be obtained by restriction from Borel measures on $\mathbb{R}^N$.

The **Borel sets** in K are the sets in $\mathcal{B}_N(K) = \mathcal{B}_N \cap K$. Using Lemma 6.24, we readily see that $\mathcal{B}_N(K)$ is the smallest σ-algebra for K containing the compact subsets of K; the argument is simpler than the corresponding proof for Proposition 6.23 in that it is not necessary to produce some sequence of sets by means of Lemma 6.22.

A **Borel function** on the compact set K is a scalar-valued function measurable with respect to $\mathcal{B}_N(K)$. A **Borel measure** on K is a measure on $\mathcal{B}_N(K)$ that is finite on compact subsets of K. In this situation regularity is a consequence of the regularity of Borel measures on $\mathbb{R}^N$, and no separate argument is needed. In fact, if μ is a Borel measure on K, we can define $\nu(E) = \mu(E \cap K)$ for each Borel subset E on $\mathbb{R}^N$, and then the finiteness of $\mu(K)$ implies that ν is a Borel measure on $\mathbb{R}^N$. Borel measures on $\mathbb{R}^N$ are regular, and therefore we have

$$\nu(E) = \sup_{\substack{K' \subseteq E,\\ K' \text{ compact in } \mathbb{R}^N}} \nu(K') = \inf_{\substack{U \supseteq E,\\ U \text{ open in } \mathbb{R}^N}} \nu(U).$$

Replacing E by $E \cap K$ and substituting from the definition of ν, we obtain the following proposition.

Proposition 6.30. Every Borel measure μ on a compact nonempty set K in $\mathbb{R}^N$ is **regular** in the sense that the value of μ on any Borel set E in K is given by

$$\mu(E) = \sup_{\substack{K \subseteq E,\\ K' \text{ compact in } K}} \mu(K') = \inf_{\substack{U \supseteq E,\\ U \text{ relatively}\\ \text{open in } K}} \mu(U).$$

4. Comparison of Riemann and Lebesgue Integrals

This section contains the definitive theorem about the relationship between the Riemann integral and the Lebesgue integral in $\mathbb{R}^N$. The Riemann integral is defined in Section III.7, the Lebesgue integral is defined in Section V.3, and Lebesgue measure in $\mathbb{R}^N$ is defined in Section 1 of the present chapter. In order to have a notational distinction between the Riemann and Lebesgue integrals, we write in this section $\mathcal{R}\int_A f\,dx$ for the Riemann integral of a bounded real-valued function on a compact geometric rectangle A, and we write $\int_A f\,dx$ for the Lebesgue integral.

Theorem 6.31. Suppose that f is a bounded real-valued function on a compact geometric rectangle A in $\mathbb{R}^N$. Then f is Riemann integrable on A if and only if f is continuous except on a Lebesgue measurable set of Lebesgue measure 0. In this case, f is Lebesgue measurable, and $\mathcal{R}\int_A f\,dx = \int_A f\,dx$.

PROOF. Proposition 6.12 shows that a set in $\mathbb{R}^N$ has "measure 0" in the sense of Chapter III if and only if it is Lebesgue measurable of measure 0, and Theorem 3.29 shows that f is Riemann integrable on A if and only if f is continuous except on a set of measure 0. This proves the first conclusion of the theorem.

For the second conclusion, suppose that $\mathcal{R}\int_A f\,dx$ exists. Lemma 3.23 shows that there exists a sequence of partitions $P^{(k)}$ of A, each refining the previous one, such that the lower Riemann sums $L(P^{(k)}, f)$ increase to $\mathcal{R}\int_A f\,dx$ and the upper Riemann sums $U(P^{(k)}, f)$ decrease to $\mathcal{R}\int_A f\,dx$. For each k, we define Borel functions L_k and U_k on A as follows: If x is an interior point of some component (closed) rectangle S of $P^{(k)}$, we define $L_k(x) = m_S(f)$, where $m_S(f)$ is the infimum of f on S; otherwise we let $L_k(x) = 0$. If we write $|S|$ for the volume of S, then the Lebesgue integral of L_k over S is given by $\int_S L_k(x)\,dx = m_S(f)|S|$. Consequently

$$\int_A L_k(x)\,dx = \sum_S m_S(f)|S| = L(P^{(k)}, f).$$

We define $U_k(x)$ similarly, using the supremum $M_S(f)$ of f on S instead of the infimum, and then

$$\int_A U_k(x) = \sum_S M_S(f)|S| = U(P^{(k)}, f).$$

Let E be the subset of points x in A such that x is in the interior of a component rectangle of $P^{(k)}$ for all k. The set $A - E$ is a Borel set of Lebesgue measure 0. Since $P^{(k+1)}$ is a refinement of $P^{(k)}$ for every k, we have $L_k(x) \le L_{k+1}(x)$ and $U_k(x) \ge U_{k+1}(x)$ for all x in E and all k. Therefore $L(x) = \lim L_k(x)$ and $U(x) = \lim U_k(x)$ exist for x in E. Since $L_k(x) \le f(x) \le U_k(x)$ for x in E, we see that

$$L(x) \le f(x) \le U(x) \qquad \text{for all } x \text{ in } E.$$

Define $L(x) = U(x) = 0$ on $A - E$. Then L and U are Borel functions with $L(x) \le U(x)$ everywhere on A. On E, we have dominated convergence, and thus

$$\int_E L(x)\,dx = \lim_k \int_E L_k(x)\,dx \qquad \text{and} \qquad \int_E U(x)\,dx = \lim_k \int_E U_k(x)\,dx.$$

The set $A - E$ has Lebesgue measure 0, and therefore these equations imply that

$$\int_A L(x)\,dx = \lim_k \int_A L_k(x)\,dx \qquad \text{and} \qquad \int_A U(x)\,dx = \lim_k \int_A U_k(x)\,dx.$$

Consequently

$$\int_A L(x)\,dx = \lim_k \int_A L_k(x)\,dx = \lim_k L(P^{(k)}, f) = \mathcal{R}\int_A f\,dx$$
$$= \lim_k U(P^{(k)}, f) = \lim_k \int_A U_k(x)\,dx = \int_A U(x)\,dx.$$

Since $L(x) \le U(x)$ on A, Corollary 5.23 shows that the set F where $L(x) = U(x)$ is a Borel set such that $A - F$ has Lebesgue measure 0. Since the inequalities $L(x) \le f(x) \le U(x)$ are valid for x in E, $f(x)$ equals $L(x)$ at least on $E \cap F$. The set $E \cap F$ is a Borel set, and L is Borel measurable; hence the restriction of f to $E \cap F$ is Borel measurable. The set $A-(E\cap F)$ is a Borel set of measure 0, and the restriction of f to this set is Lebesgue measurable no matter what values f assumes on this set. Thus f is Lebesgue measurable. Then the Lebesgue integral $\int_A f\,dx$ is defined, and we have

$$\int_A f(x)\,dx = \int_{E\cap F} f(x)\,dx = \int_{E\cap F} L(x)\,dx = \int_A L(x)\,dx = \mathcal{R}\int_A f\,dx.$$

5. Change of Variables for the Lebesgue Integral

A general-looking change-of-variables formula for the Riemann integral was proved in Section III.10. On closer examination of the theorem, we found that the result did not fully handle even as ostensibly simple a case as the change from Cartesian coordinates in $\mathbb{R}^2$ to polar coordinates. Lebesgue integration gives us methods that deal with all the unpleasantness that was concealed by the earlier formula.

Theorem 6.32 (change-of-variables formula). Let φ be a one-one function of class C^1 from an open subset U of $\mathbb{R}^N$ onto an open subset $\varphi(U)$ of $\mathbb{R}^N$ such that $\det \varphi'(x)$ is nowhere 0. Then

$$\int_{\varphi(U)} f(y)\,dy = \int_U f(\varphi(x))|\det \varphi'(x)|\,dx$$

for every nonnegative Borel function f defined on $\varphi(U)$.

REMARK. The σ-algebra on $\varphi(U)$ is understood to be $\mathcal{B}_N \cap \varphi(U)$, the set of intersections of Borel sets in $\mathbb{R}^N$ with the open set $\varphi(U)$. If f is extended from $\varphi(U)$ to $\mathbb{R}^N$ by defining it to be 0 off $\varphi(U)$, then measurability of f with respect to this σ-algebra is the same as measurability of the extended function with respect to $\mathcal{B}_N$.

PROOF. Theorem 3.34 gives us the change-of-variables formula, as an equality of Riemann integrals, for every f in $C_{\text{com}}(\varphi(U))$. In this case the integrands on both sides, when extended to be 0 outside the regions of integration, are continuous on all of $\mathbb{R}^N$, and the integrations can be viewed as involving continuous functions on compact geometric rectangles. Proposition 6.11 (or Theorem 6.31 if one prefers) allows us to reinterpret the equality as an equality of Lebesgue integrals.

In the extension of this identity to all nonnegative Borel functions, measurability will not be an issue. The function f is to be measurable with respect to $\mathcal{B}_N(\varphi(U))$, and Corollary 6.29 shows that such f's correspond exactly to functions $f \circ \varphi$ measurable with respect to $\mathcal{B}_N(U)$.

Using Theorem 5.19, define a measure μ on $\mathcal{B}_N(U)$ by

$$\mu(E) = \int_E |\det \varphi'(x)|\, dx.$$

Corollary 5.28 implies that μ satisfies

$$\int_U g\, d\mu = \int_U g(x)\, |\det \varphi'(x)|\, dx \tag{$*$}$$

for every nonnegative g on U measurable with respect to $\mathcal{B}_N(U)$. Next define another set function ν on $\mathcal{B}_N(U)$ by

$$\nu(E) = m(\varphi(E)),$$

where m is Lebesgue measure. It is immediate that ν is a measure, and we have $\int_{\varphi(U)} I_E(\varphi^{-1}(y))\, dy = \int_{\varphi(U)} I_{\varphi(E)}\, dy = m(\varphi(E)) = \nu(E) = \int_U I_E\, d\nu$. Passing to simple functions ≥ 0 and then using monotone convergence, we obtain

$$\int_{\varphi(U)} g \circ \varphi^{-1}\, dy = \int_U g\, d\nu \tag{$**$}$$

for every nonnegative g on U measurable with respect to $\mathcal{B}_N(U)$.

If in ($**$) and ($*$) we take $g = f \circ \varphi$ with f in $C_{\text{com}}(\varphi(U))$ and we substitute into the change-of-variables formula as it is given for f in $C_{\text{com}}(\varphi(U))$, we obtain the identity

$$\int_U g\, d\nu = \int_U g\, d\mu \tag{$\dagger$}$$

for all g in $C_{\text{com}}(U)$. From Corollary 6.26 we conclude that $\mu = \nu$. Hence ($\dagger$) holds for every nonnegative g on U measurable with respect to $\mathcal{B}_N(U)$. We unwind ($\dagger$) using ($**$) and ($*$) with $g = f \circ \varphi$ but now taking f to be any nonnegative function on $\varphi(U)$ measurable with respect to $\mathcal{B}_N(\varphi(U))$, and we obtain the conclusion of the theorem.

Let us return to the example of polar coordinates in $\mathbb{R}^2$, first considered in Section III.10. The data in the theorem are

$$U = \left\{ \begin{pmatrix} r \\ \theta \end{pmatrix} \;\Big|\; 0 < r < +\infty \text{ and } 0 < \theta < 2\pi \right\},$$
$$\begin{pmatrix} x \\ y \end{pmatrix} = \varphi \begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix},$$

and we have

$$\varphi(U) = \mathbb{R}^2 - \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \;\Big|\; x \geq 0 \right\}.$$

Since $\det\varphi'\begin{pmatrix} r \\ \theta \end{pmatrix} = r$, Theorem 6.32 gives

$$\int_{\varphi(U)} f(x, y)\, dx\, dy = \int_{0<r<\infty,\ 0<\theta<2\pi} f(r\cos\theta, r\sin\theta)\, r\, dr\, d\theta$$

for every nonnegative Borel function f on $\varphi(U)$. The set of integration $\varphi(U)$ on the left side is not quite the whole plane; it omits the part of the x axis where $x \geq 0$. But this is a harmless defect: this subset of the x axis is contained in the entire x axis, which is an abstract rectangle in the sense of Fubini's Theorem and has measure 0. Thus the formula can be changed to read

$$\int_{\mathbb{R}^2} f(x, y)\, dx\, dy = \int_{0\leq r<\infty,\ 0\leq\theta<2\pi} f(r\cos\theta, r\sin\theta)\, r\, dr\, d\theta$$

for every nonnegative Borel function f on $\mathbb{R}^2$. Here is an application of this formula that we shall use in proving the Fourier Inversion Formula in Chapter VIII.

Proposition 6.33. $\displaystyle\int_{-\infty}^{\infty} e^{-\pi x^2}\, dx = 1.$

REMARK. Since we now know from Theorem 6.31 that there is no discrepancy between the Riemann integral and the Lebesgue integral with respect to Lebesgue measure, there will be no harm in the future in writing limits of integration in the usual way for integrals with respect to 1-dimensional Lebesgue measure.

PROOF. We use polar coordinates and Fubini's Theorem to compute the square of the integral in question:

$$\begin{aligned}
\left(\int_{\mathbb{R}} e^{-\pi x^2}\, dx\right)^2 &= \int_{\mathbb{R}^2} e^{-\pi x^2} e^{-\pi y^2}\, dx\, dy = \int_{\mathbb{R}^2} e^{-\pi(x^2+y^2)}\, dx\, dy \\
&= \int_0^\infty \int_0^{2\pi} e^{-\pi r^2} r\, d\theta\, dr = 2\pi \int_0^\infty r e^{-\pi r^2}\, dr \\
&= 2\pi \lim_N \int_0^N r e^{-\pi r^2}\, dr = \lim_N \left[- e^{-\pi r^2} \right]_0^N = 1.
\end{aligned}$$

Since the integral in question is certainly > 0, the proposition follows.

Proposition 6.33 is closely related to properties of the "gamma function," a certain function of a complex variable that reduces essentially to the factorial function on the positive integers. The definition of the gamma function makes use of the expression t^s defined for $0 < t < +\infty$ and s in $\mathbb{C}$ by $t^s = e^{s \log t}$.

Fix $s \in \mathbb{C}$ with $\operatorname{Re} s > 0$. The function $t \mapsto t^{s-1}e^{-t}$ is continuous on $(0, +\infty)$ and hence Borel measurable. Let us see that it is integrable with respect to Lebesgue measure. Since $|t^{s-1}e^{-t}| = t^{\operatorname{Re} s - 1}e^{-t}$, we may assume that s is real (and positive) in showing the integrability. Integrability on $(0, 1]$ is no problem, since we know that $\int_0^1 t^{s-1}\,dt < \infty$ for $s > 0$. To handle $[1, +\infty)$, let n be an integer $\geq s - 1$. Then $t^{s-1} \leq t^n = 2^n n!\big(\frac{1}{n!}\big(\frac{t}{2}\big)^n\big) \leq 2^n n! \sum_{k=0}^\infty \frac{1}{k!}\big(\frac{t}{2}\big)^k = 2^n n! e^{t/2}$. Hence $t^{s-1}e^{-t} \leq 2^n n! e^{-t/2}$, and the integrability on $[1, +\infty)$ follows.

With this integrability in place, we define the **gamma function** by

$$\Gamma(s) = \int_0^\infty t^{s-1}e^{-t}\,dt \qquad \text{for } \operatorname{Re} s > 0.$$

Proposition 6.34. The gamma function has the properties that

(a) $\Gamma(s + 1) = s\Gamma(s)$ for $\operatorname{Re} s > 0$,
(b) $\Gamma(1) = 1$ and $\Gamma(n + 1) = n!$ for integers $n > 0$,
(c) $\Gamma\big(\frac{1}{2}\big) = \sqrt{\pi}$.

PROOF. Part (a) follows from integration by parts, which needs to be done on an interval $[\varepsilon, M]$ and followed by passages to the limit $\varepsilon \to 0$ and $M \to \infty$. In (b), the formula $\Gamma(1) = 1$ just amounts to the elementary integral $\int_0^\infty e^{-t}\,dt = 1$, and then the formula $\Gamma(n + 1) = n!$ for integers $n > 0$ follows by iterating (a). For (c), the change of variables $t = \pi x^2$ gives

$$\Gamma\big(\tfrac{1}{2}\big) = \int_0^\infty t^{-1/2}e^{-t}\,dt = \int_0^\infty (\pi x^2)^{-1/2}e^{-\pi x^2}2\pi x\,dx = 2\sqrt{\pi}\int_0^\infty e^{-\pi x^2}\,dx.$$

Since $\int_0^\infty e^{-\pi x^2}\,dx = \frac{1}{2}\int_{-\infty}^\infty e^{-\pi x^2}\,dx$, Proposition 6.33 allows us to conclude that $2\int_0^\infty e^{-\pi x^2}\,dx = 1$. Hence $\Gamma\big(\frac{1}{2}\big) = \sqrt{\pi}$.

It is often true in applications of the change-of-variables formula that the set $\varphi(U)$ does not exhaust the set that one might hope to have as region of integration. For polar coordinates the exceptional set was the part of the x axis with $x \geq 0$, and an easy argument showed that the exceptional set had measure 0. In a more complicated example, that easy argument will not ordinarily apply, but still the exceptional set has a certain "lower-dimensional" quality to it. A general result saying that certain lower-dimensional sets have measure 0 will be given as a corollary of Sard's Theorem, which we prove now.

Let $\psi : V \to \mathbb{R}^N$ be a smooth map defined on an open subset V of $\mathbb{R}^N$. A **critical point** x of ψ is a point where $\psi'(x)$ has rank $< N$. In this case, $\psi(x)$ is called a **critical value**.

Theorem 6.35 (Sard's Theorem). If $\psi : V \to \mathbb{R}^N$ is a smooth map defined on an open subset V of $\mathbb{R}^N$, then the set of critical values of ψ is a Borel set of Lebesgue measure 0 in $\mathbb{R}^N$.

PROOF. The set where $\psi'(x)$ has rank $\le N-1$ is relatively closed in V and hence is the union of countably many compact sets. The set of critical values is then the union of the compact images of these sets and consequently is a Borel set. Let us see that this Borel set has Lebesgue measure 0. Since V is the countable union of compact geometric rectangles and since the countable union of sets of measure 0 is of measure 0, it is enough to prove the theorem for the restriction of ψ to a compact geometric rectangle R.

For points $x = (x_1, \dots, x_N)$ and $x' = (x'_1, \dots, x'_N)$ in R, the Mean Value Theorem gives

$$\psi_i(x') - \psi_i(x) = \sum_{j=1}^{N} \frac{\partial \psi_i}{\partial x_j}(z_i)(x'_j - x_j), \tag{$*$}$$

where z_i is a point on the line segment from x to x'. Since the $\frac{\partial \psi_i}{\partial x_j}$ are bounded on R, we see as a consequence that

$$|\psi(x') - \psi(x)| \le a|x' - x| \tag{$**$}$$

with a independent of x and x'. Let $L_x(x') = (L_{x,1}(x'), \dots, L_{x,N}(x'))$ be the best first-order approximation to ψ about x, namely

$$L_{x,i}(x') = \psi_i(x) + \sum_{j=1}^{N} \frac{\partial \psi_i}{\partial x_j}(x)(x'_j - x_j).$$

Subtracting this equation from $(*)$, we obtain

$$\psi_i(x') - L_{x,i}(x') = \sum_{j=1}^{N} \left(\frac{\partial \psi_i}{\partial x_j}(z_i) - \frac{\partial \psi_i}{\partial x_j}(x) \right) (x'_j - x_j).$$

Since $\frac{\partial \psi_i}{\partial x_j}$ is smooth and $|z_i - x| \le |x' - x|$, we deduce that

$$|\psi(x') - L_x(x')| \le b|x' - x|^2 \tag{$\dagger$}$$

with b independent of x and x'.

If x is a critical point, let us bound the image of the set of x' with $|x' - x| \le c$. The determinant of the linear part of L_x is 0, and hence L_x has image in a hyperplane, not necessarily a coordinate hyperplane. By $(\dagger)$, $\psi(x')$ has distance

$\leq bc^2$ from this hyperplane. In each of the $N-1$ perpendicular directions, $(**)$ shows that $\psi(x')$ and $\psi(x)$ are at distance $\leq ac$ from each other. Thus $\psi(x')$ is contained in a box[1] centered at $\psi(x)$ with volume $2^N(ac)^{N-1}(bc^2) = 2^N a^{N-1} b c^{N+1}$.

We subdivide R into M^N smaller compact geometric rectangles whose dimensions are $1/M$ times those of R. If d is the diameter of R and if one of these smaller geometric rectangles R' contains a critical point x, then any point x' in R' has $|x'-x| \leq d/M$. By the result of the previous paragraph, ψ of R' is contained in a box of volume $2^N a^{N-1} b(d/M)^{N+1}$. The union of these boxes, taken over all of the smaller geometric rectangles containing critical points, contains the critical values. Since there are at most M^N of the smaller geometric rectangles, the outer measure m^* of the set of critical values, where m refers to Lebesgue measure, is $\leq 2^N a^{N-1} b d^{N+1} M^{-1}$. This estimate is valid for all M, and hence the set S of critical values has $m^*(S) = 0$. Therefore the Borel set S has Lebesgue measure 0.

Corollary 6.36. If $\psi : V \to \mathbb{R}^N$ is a smooth map defined on an open subset V of $\mathbb{R}^M$ with $M < N$, then the image of ψ is a Borel set of Lebesgue measure 0 in $\mathbb{R}^N$.

PROOF. Sard's Theorem (Theorem 6.35) applies to the composition of the projection $\mathbb{R}^N \to \mathbb{R}^M$ followed by ψ. Every point of the domain is a critical point, and hence every point of the image is a critical value. The result follows.

We define a **lower-dimensional set** in $\mathbb{R}^N$ to be any set contained in the countable union of smooth images of open sets in Euclidean spaces of dimension $< N$. The following result is immediate from Corollary 6.36.

Corollary 6.37. Any lower-dimensional set in $\mathbb{R}^N$ is Lebesgue measurable of Lebesgue measure 0.

The N-dimensional generalization of polar coordinates in $\mathbb{R}^2$ is **spherical coordinates** in $\mathbb{R}^N$. In the notation of Theorem 6.32, we have

$$U = \left\{ \begin{pmatrix} r \\ \theta_1 \\ \vdots \\ \theta_{N-1} \end{pmatrix} \;\middle|\; \begin{array}{l} 0 < r < +\infty, \\ 0 < \theta_j < \pi \text{ for } 1 \leq j \leq N-2, \\ 0 < \theta_{N-1} < 2\pi \end{array} \right\}$$

and
$$\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \varphi \begin{pmatrix} r \\ \theta_1 \\ \vdots \\ \theta_{N-1} \end{pmatrix} = \begin{pmatrix} r\cos\theta_1 \\ r\sin\theta_1\cos\theta_2 \\ \vdots \\ r\sin\theta_1\cdots\sin\theta_{N-2}\cos\theta_{N-1} \\ r\sin\theta_1\cdots\sin\theta_{N-2}\sin\theta_{N-1} \end{pmatrix}.$$

[1]This box need not have its faces parallel to the coordinate hyperplanes.

Problem 2 at the end of the chapter asks for three things to be checked:

(i) the determinant factor in the change-of-variables formula is given by

$$|\det \varphi'| = r^{N-1} \sin^{N-2}\theta_1 \sin^{N-3}\theta_2 \cdots \sin\theta_{N-2},$$

(ii) φ is one-one on U,

(iii) the complement of $\varphi(U)$ in $\mathbb{R}^N$ is a lower-dimensional set.

Then it follows that the change-of-variables formula applies and that the integration over $\varphi(U)$ can be extended over $\mathbb{R}^N$. We can write the result as

$$\int_{\mathbb{R}^N} f(x)\,dx = \int_{r=0}^{\infty}\int_{\theta_1=0}^{\pi}\cdots\int_{\theta_{N-2}=0}^{\pi}\int_{\theta_{N-1}=0}^{2\pi} f(r\cos\theta_1,\dots)$$
$$\times\, r^{N-1}\sin^{N-2}\theta_1\cdots\sin\theta_{N-2}\,d\theta_{N-1}\cdots d\theta_1\,dr.$$

The expression $\sin^{N-2}\theta_1\cdots\sin\theta_{N-2}\,d\theta_{N-1}\cdots d\theta_1$ we abbreviate as $d\omega$. Geometrically it is the contribution to Lebesgue measure on $\mathbb{R}^N$ from the sphere S^{N-1} of radius 1 centered at the origin. In Chapter XI we shall speak of Borel sets in any compact metric space. The sphere S^{N-1} is a compact metric space, and we shall note that $d\omega$ refers to a rotation-invariant Borel measure on S^{N-1}. We write

$$\Omega_{N-1} = \int_{S^{N-1}} d\omega$$

for the "area" of the sphere S^{N-1}. This constant is evaluated in Problem 12 at the end of the present chapter with the aid of Proposition 6.33. In terms of $d\omega$, the change-of-variables formula for spherical coordinates is

$$\int_{\mathbb{R}^N} f(x)\,dx = \int_{r=0}^{\infty}\int_{\omega\in S^{N-1}} f(r\omega)\,r^{N-1}\,d\omega\,dr.$$

This formula allows us quickly to check the integrability of powers of $|x|$ near the origin and near ∞. In fact, we have

$$\int_{|x|\le 1} |x|^q\,dx = \Omega_{N-1}\int_0^1 r^{q+N-1}\,dr$$

and

$$\int_{|x|\ge 1} |x|^q\,dx = \Omega_{N-1}\int_1^{\infty} r^{q+N-1}\,dr,$$

from which we see that

$$|x|^q \text{ is integrable near } \begin{cases} 0 & \text{for } q > -N, \\ \infty & \text{for } q < -N. \end{cases}$$

6. Hardy–Littlewood Maximal Theorem

This section takes a first look at the theory of almost-everywhere convergence. The theory developed historically out of Lebesgue's work on an extension of the Fundamental Theorem of Calculus to general integrable functions on intervals of the line, work that we address largely in the next chapter. We shall see gradually that the theory applies to a broader range of problems than the ones immediately generalizing Lebesgue's work, and one can make a case that nowadays the theory in this section is of considerably greater significance in real analysis than one might expect from Lebesgue's work on the Fundamental Theorem.

The theory brings together two threads. The first thread is the observation that an effort to differentiate integrals of general integrable functions on an interval of the line can be reinterpreted as a problem of almost-everywhere convergence in connection with an approximate identity of the kind in Theorem 6.20. In explaining this assertion, let us denote Lebesgue measure by m as necessary. To differentiate $F(x) = \int_a^x f(t)\,dt$, one forms the usual difference quotient $h^{-1}[F(x+h) - F(x)]$, which can be written for $h > 0$ as

$$\frac{1}{m([-h,0])}\int_{[-h,0]} f(x-y)\,dy = \int_{\mathbb{R}^1} f(x-y)\,m([-h,0])^{-1} I_{[-h,0]}(y)\,dy$$

or as $f * \varphi_h(x)$, where $\varphi(y) = m([-1,0])^{-1} I_{[-1,0]}(y)$. Here φ has integral 1, and φ_h is the normalized dilated function defined in Section 2 by $\varphi_h(y) = h^{-1}\varphi(h^{-1}y)$ in the 1-dimensional case. Theorem 6.20 says for $p = 1$ and $p = 2$ that as h decreases to 0, $f * \varphi_h$ converges to f in L^p if f is in L^p. Also, $f * \varphi_h$ converges uniformly to f if f is bounded and uniformly continuous, and $f * \varphi_h(x)$ converges to $f(x)$ at the point x if f is bounded and is continuous at x. The problem about differentiation of integrals asks about convergence almost everywhere.

We shall want to have a theorem in $\mathbb{R}^N$, and for this purpose an N-dimensional version of $I_{[-1,0]}$ does not seem attractive for generalizing. Instead, let us generalize from $I_{[-1,1]}$, taking the N-dimensional problem to involve a ball B of radius 1 centered at the origin; there is some flexibility in choosing the set B, and a cube centered at the origin would work as well. We write rB for the set of dilates of the members of B by the scalar r. Thus we investigate

$$m(rB)^{-1}\int_{rB} f(x-y)\,dy$$

as r decreases to 0; equivalently we investigate

$$f * \varphi_r(x), \quad \text{where } \varphi(y) = m(B)^{-1} I_B(y).$$

The second thread comes from making a simple observation and then trying to prove the converse in specific settings, as improbable as it sounds. The observation is that if some sequence of nonnegative functions indexed by n is to converge almost everywhere, its supremum on n must be finite almost everywhere. A converse would say that a finite supremum almost everywhere implies convergence almost everywhere. Banach succeeded in proving an abstract such converse in a 1926 paper, making use of the completeness of the space of functions he was studying. In a celebrated 1930 paper, Hardy and Littlewood proved a concrete such converse in connection with differentiation of integrals; they obtained a quantitative estimate about the supremum, and then the almost everywhere convergence followed from that estimate and from the fact that the convergence certainly takes place for nice functions. Here is an N-dimensional version of the basic theorem in that direction.

Theorem 6.38 (Hardy–Littlewood Maximal Theorem). If f is in $L^1(\mathbb{R}^N)$, then

$$m\{x \mid f^*(x) > \xi\} \le \frac{5^N \|f\|_1}{\xi}$$

for every $\xi > 0$, where

$$f^*(x) = \sup_{0<r<\infty} m(rB)^{-1} \int_{rB} |f(x-y)|\,dy.$$

Before examining the statement of the theorem more closely and then proving the theorem, let us see how to derive a corresponding N-dimensional convergence result from it, and let us see how the first part of Lebesgue's version of the Fundamental Theorem of Calculus, the part about differentiation of integrals, follows as well.

Corollary 6.39. If f is integrable on every bounded subset of $\mathbb{R}^N$, then

$$\lim_{r\downarrow 0} m(rB)^{-1} \int_{rB} f(x-y)\,dy = f(x) \quad \text{a.e.}$$

PROOF. Since the convergence for a particular x depends on the behavior of the function only near x, we may assume that f is identically 0 off some bounded set. The effect of this assumption for our purposes is that f then has to be in $L^1(\mathbb{R}^N)$. Define

$$T_r(f) = m(rB)^{-1} \int_{rB} f(x-y)\,dy,$$

bearing in mind that $f^*(x) = \sup_{r>0} T_r(|f|)(x)$. If g is continuous of compact support, then $\lim_{r\downarrow 0} T_r g(x) = g(x)$ everywhere by Theorem 6.20c. Let $\epsilon > 0$ be given, and choose by Corollary 6.4 a function g in $C_{\mathrm{com}}(\mathbb{R}^N)$ with $\|f-g\|_1 < \epsilon$. Then

$$\begin{aligned}
&\limsup_{r\downarrow 0} |T_r f(x) - f(x)| \\
&\qquad \le \limsup_{r\downarrow 0} |T_r(f-g)(x)| + \limsup_{r\downarrow 0} |T_r g(x) - g(x)| + |g(x) - f(x)| \\
&\qquad \le \sup_{r>0} |T_r(f-g)(x)| + |g(x) - f(x)| \\
&\qquad \le \sup_{r>0} T_r(|f-g|)(x) + |g(x) - f(x)|.
\end{aligned}$$

If the left side is $> \xi$, at least one of the terms on the right side is $> \xi/2$. Hence

$$\begin{aligned}
&\Big\{x \;\Big|\; \limsup |T_r f(x) - f(x)| > \xi\Big\} \\
&\qquad\qquad \subseteq \Big\{x \;\Big|\; (f-g)^*(x) > \xi/2\Big\} \cup \Big\{x \;\Big|\; |f(x) - g(x)| > \xi/2\Big\}.
\end{aligned}$$

By Theorem 6.38 and the inequality $\alpha\, m\big\{x \;\big|\; |F(x)| > \alpha\big\} \le \|F\|_1$, the Lebesgue measure of the right side is

$$\le \frac{2 \cdot 5^N \|f-g\|_1}{\xi} + \frac{2\|f-g\|_1}{\xi} \le \epsilon\, \frac{2(5^N+1)}{\xi}.$$

Since ϵ is arbitrary, $S(\xi) = \big\{x \;\big|\; \limsup |T_r f(x) - f(x)| > \xi\big\}$ has measure 0. Letting ξ tend to 0 through the values $1/n$, we see that $S(0)$ has measure 0, i.e., that $\lim_{r\downarrow 0} T_r f(x) = f(x)$ almost everywhere.

Corollary 6.40 (first part of Lebesgue's form of the Fundamental Theorem of Calculus). If f is integrable on every bounded subset of $\mathbb{R}^1$, then $\int_a^x f(y)\,dy$ is differentiable almost everywhere and

$$\frac{d}{dx}\int_a^x f(y)\,dy = f(x) \quad \text{a.e.}$$

PROOF. For f in $L^1(\mathbb{R}^1)$, let f^* be as in Theorem 6.38, and define

$$f_r^{**}(x) = \sup_{h>0} \frac{1}{h}\int_0^h |f(x+t)|\,dt \quad \text{and} \quad f_l^{**}(x) = \sup_{h>0} \frac{1}{h}\int_{-h}^0 |f(x+t)|\,dt.$$

Then

$$f_r^{**}(x) \le \sup_{h>0} \frac{1}{h}\int_{-h}^h |f(x+t)|\,dt = 2f^*(x),$$

and similarly $f_l^{**}(x) \le 2f^*(x)$. From Theorem 6.38 it follows that

$$m\big\{x \bigm| f_r^{**}(x) > \xi\big\} \le 10\|f\|_1\big/\xi$$

and

$$m\big\{x \bigm| f_l^{**}(x) > \xi\big\} \le 10\|f\|_1\big/\xi.$$

The same argument as for Corollary 6.39 allows us to conclude, for any f integrable on every bounded subset of $\mathbb{R}^1$, that $\lim_{h\downarrow 0} \frac{1}{h}\int_0^h f(x+t)\,dt = f(x)$ a.e. and $\lim_{h\downarrow 0} \frac{1}{h}\int_{-h}^0 f(x+t)\,dt = f(x)$ a.e. Hence $\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$ almost everywhere for such f.

Let us return to Theorem 6.38. The function $f^*(x)$ is called the **Hardy–Littlewood maximal function** of f. It is measurable because the supremum over rational r gives the same value of $f^*(x)$ for each x. If we let ξ tend to ∞ in the inequality $m\big\{x \bigm| f^*(x) > \xi\big\} \le 5^N\big\|f\big\|_1\big/\xi$, we see immediately that $f^*(x)$ is finite almost everywhere, i.e., that the supremum in question is actually finite almost everywhere. The inequality is a quantitative version of that qualitative conclusion.

For any situation in which it is desired to prove an almost-everywhere convergence theorem, there is an associated **maximal function** in modern terminology, which can be taken as the supremum of the absolute value of the quantity for which one is trying to prove almost-everywhere convergence. In the above case we used the supremum for $|f|$ instead, which in principle could be larger.

There is no hope that the Hardy–Littlewood maximal function f^* is actually in L^1 if f is not a.e. the 0 function because the occurrence of large values of r in the supremum already rules out L^1 behavior: in fact, $f^*(x)$ is necessarily $\ge$ a positive multiple of $|x|^{-N}$ for large $|x|$, and thus f^* cannot be integrable. On the other hand, f^* is close to integrable: We shall see in Section 10 that the integral of any nonnegative function g can be computed in terms of the function $m\big\{x \bigm| g(x) > \xi\big\}$ of ξ, the formula being $\int_{\mathbb{R}^N} g(x)\,dx = \int_0^\infty m\big\{x \bigm| g(x) > \xi\big\}\,d\xi$. Theorem 6.38 shows that the integrand in the case of f^* is $\le$ a multiple of $1/\xi$, and $1/\xi$ is close to being integrable on $(0, +\infty)$. This is a better qualitative conclusion than merely finiteness almost everywhere, and Theorem 6.38 is a quantitative version of just how close f^* is to being integrable.

The particular property of f^* that is isolated in Theorem 6.38 arises fairly often. If $g \ge 0$ is integrable and S is the set where $g > \xi$, then $g \ge \xi I_S$ everywhere; hence $\|g\|_1 \ge \xi\, m(S)$ and $m(S) \le \|g\|_1\big/\xi$. A function g is said to be in **weak** L^1 if

$$m\big\{x \bigm| |g(x)| > \xi\big\} \le C\big/\xi$$

for some constant C and for all $\xi > 0$. Theorem 6.38 says that the nonlinear operator $f \mapsto f^*$ carries L^1 to weak L^1 with C bounded by a multiple of the L^1

norm of f, and an operator of this kind that satisfies also a certain sublinearity property is said to be of **weak type** $(1, 1)$. We return to this matter, with the definition in a clearer context, in Chapter IX.

Now let us prove Theorem 6.38. One modern proof uses the following covering lemma, which takes into account the geometry of $\mathbb{R}^N$ in a surprisingly subtle way. Once the lemma is in hand, the rest is easy.

Lemma 6.41 (Wiener's Covering Lemma). Let $E \subseteq \mathbb{R}^N$ be a Borel set, and suppose that to each x in E there is associated some ball $B(r; x)$ with r perhaps depending on x. If the radii $r = r(x)$ are bounded, then there is a finite or countable *disjoint* collection of these balls, say $B(r_1; x_1), B(r_2; x_2), \ldots$, such that either the collection is infinite and $\inf_{1 \leq j < \infty} r_j \neq 0$ or

$$E \subseteq \bigcup_{j=1}^{\infty} B(5r_j; x_j).$$

In either case,

$$m(E) \leq 5^N \sum_{j=1}^{\infty} m(B(r_j; x_j)).$$

REMARK. The shape of the sets of $B(r; x)$ is not very important. What is important is that there be some neighborhood B of the origin that is closed under the operation of multiplying all its members by -1 and by any positive number $r \leq 1$. The other sets are obtained from B by dilation and translation.

PROOF. Let

$$\mathcal{A}_1 = \big\{\text{all sets } B(r; x) \text{ in question}\big\}$$

and

$$R_1 = \sup \big\{r \ \big|\ B(r; x) \text{ is in } \mathcal{A}_1 \text{ for some } x\big\}.$$

By hypothesis, R_1 is finite. Pick some $B(r_1; x_1)$ with $r_1 \geq \frac{1}{2} R_1$, and let

$$\mathcal{A}_2 = \big\{\text{members of } \mathcal{A}_1 \text{ disjoint from } B(r_1; x_1)\big\}.$$

If $\mathcal{A}_2$ is empty, let all further R_j's be 0 and let all further $B(r_j; x_j)$'s be empty. If $\mathcal{A}_2$ is nonempty, let

$$R_2 = \sup \big\{r \ \big|\ B(r; x) \text{ is in } \mathcal{A}_2 \text{ for some } x\big\}.$$

Pick $B(r_2; x_2)$ in $\mathcal{A}_2$ with $r_2 \geq \frac{1}{2} R_2$. Let

$$\mathcal{A}_3 = \big\{\text{members of } \mathcal{A}_2 \text{ disjoint from } B(r_2; x_2)\big\},$$

and proceed inductively to construct R_3, $B(r_3; x_3)$, $\mathcal{A}_4$, etc.

The numbers R_j are monotone decreasing. We may assume that $\lim R_j = 0$, since otherwise $\inf_j r_j \neq 0$ and $\sum m(B(r_j; x_j)) = +\infty$. Let

$$V_j = \text{union of all sets in } \mathcal{A}_j - \mathcal{A}_{j+1} \qquad \text{for } j \geq 1$$

and
$$V_0 = \text{union of all sets in } \mathcal{A}_1.$$

Then $V_0 = \bigcup_{j=1}^{\infty} V_j$; in fact, if $B(r; x)$ is in $\mathcal{A}_1$, then the equality $\lim R_j = 0$ forces there to be a last index j such that $B(r; x)$ is in $\mathcal{A}_j$, and this j has the property that $B(r; x)$ is in $\mathcal{A}_j$ and not $\mathcal{A}_{j+1}$.

Since $E \subseteq \bigcup_{x \in E} B(r; x) = V_0 = \bigcup_{j=1}^{\infty} V_j$, the proof will be complete if we show that

$$V_j \subseteq B(5r_j; x_j). \qquad (*)$$

Thus let $B(r; x)$ be in $\mathcal{A}_j - \mathcal{A}_{j+1}$. Then $r \leq R_j$,

$$B(r; x) \cap B(r_j; x_j) \neq \varnothing,$$

and $r_j \geq \frac{1}{2} R_j$. Consequently $r \leq 2r_j$ and

$$B(r; x - x_j) \cap B(r_j; 0) \neq \varnothing.$$

This condition means that there is some p in $B(r_j; 0)$ with $|x - x_j - p| < r$. If q is any member of $B(r; x)$, then

$$|q - x_j| \leq |q - x| + |x - x_j - p| + |p| < r + r + r_j = 2r + r_j.$$

Thus q is in $B(2r + r_j; x_j) \subseteq B(5r_j; x_j)$, and $(*)$ follows.

PROOF OF THEOREM 6.38. Let $E = \{x \mid f^*(x) > \xi\}$. If x is in E, then $m(B(r; 0))^{-1} \int_{B(r;x)} |f(y)|\,dy > \xi$ for some $r > 0$. Associate this r to x in applying Lemma 6.41. Since

$$\xi < m(B(r; 0))^{-1} \int_{B(r;x)} |f(y)|\,dy \leq r^{-N} m(B(1; 0))^{-1} \|f\|_1,$$

we see that $r^N \leq \xi^{-1} m(B(1; 0))^{-1} \|f\|_1$. Hence the numbers r are bounded. Thus the lemma applies, and we obtain

$$m(E) \leq 5^N \sum_j m(B(r_j; x_j)) \leq 5^N \xi^{-1} \sum_j \int_{B(r_j; x_j)} |f(y)|\,dy \leq 5^N \xi^{-1} \|f\|_1,$$

the last inequality holding because of the disjointness of the sets $B(r_j; x_j)$.

Let us return to the theme of almost-everywhere convergence in connection with approximate identities. Theorem 6.38 has the following consequence of just that kind.

Corollary 6.42. Let $\varphi \ge 0$ be a continuous integrable function on $\mathbb{R}^N$ of the form $\varphi(x) = \varphi_0(|x|)$, where φ_0 is a decreasing C^1 function on $[0, \infty)$, and define $\varphi_\varepsilon(x) = \varepsilon^{-N}\varphi(\varepsilon^{-1}x)$ for $\varepsilon > 0$. Then there is a constant C_φ such that

$$\sup_{\varepsilon>0} |(\varphi_\varepsilon * f)(x)| \le C_\varphi f^*(x)$$

for all x in $\mathbb{R}^N$ and for all f in $L^1(\mathbb{R}^N)$. Consequently if $\int_{\mathbb{R}^N} \varphi(x)\,dx = 1$, then

$$\lim_{\varepsilon\downarrow 0} (\varphi_\varepsilon * f)(x) = f(x)$$

almost everywhere for each f in $L^1(\mathbb{R}^N)$.

PROOF. Put $\psi(r) = -\varphi_0'(r) \ge 0$, so that $\varphi_0(r) - \varphi_0(R) = \int_r^R \psi(s)\,ds$. The integrability of φ and the fact that φ_0 is decreasing force $\lim_{R\to\infty} \varphi_0(R) = 0$, and we obtain $\varphi_0(r) = \int_r^\infty \psi(s)\,ds$ and $\varphi(x) = \int_{|x|}^\infty \psi(r)\,dr$. Meanwhile, the integrability of φ, together with the formula for integrating in spherical coordinates, shows that $\int_0^\infty \varphi_0(r)\,r^{N-1}\,dr = C < +\infty$. Integrating by parts on the interval $[0, M]$ gives

$$C \ge \int_0^M \varphi_0(r)\,r^{N-1}\,dr = \tfrac{1}{N}\big[\varphi_0(r)\,r^N\big]_0^M + \tfrac{1}{N}\int_0^M \psi(r)\,r^N\,dr,$$

and thus

$$\tfrac{1}{N}\int_0^\infty \psi(r)\,r^N\,dr \le C < +\infty.$$

The form of φ implies that

$$\varphi_\varepsilon(x) = \varepsilon^{-N}\int_{\varepsilon^{-1}|x|}^\infty \psi(r)\,dr.$$

If, as in the statement of Theorem 6.38, we let B be the ball of radius 1 centered at the origin, we obtain

$$\begin{aligned}
|(\varphi_\varepsilon * f)(x)| &\le \int_{\mathbb{R}^N} \varphi_\varepsilon(y)|f(x-y)|\,dy \\
&= \int_{y\in\mathbb{R}^N} \varepsilon^{-N}\int_{r=\varepsilon^{-1}|y|}^\infty \psi(r)|f(x-y)|\,dr\,dy \\
&= \int_{r=0}^\infty \psi(r)\big[\varepsilon^{-N}\int_{|y|\le\varepsilon r} |f(x-y)|\,dy\big]\,dr \\
&= \int_{r=0}^\infty m(B)\psi(r)\,r^N\big[m(\varepsilon r B)^{-1}\int_{|y|\le\varepsilon r} |f(x-y)|\,dy\big]\,dr \\
&\le m(B)\big[\int_{r=0}^\infty \psi(r)\,r^N\,dr\big]f^*(x).
\end{aligned}$$

The right side is $\le C_\varphi f^*(x)$ with $C_\varphi = CNm(B)$. Applying Theorem 6.38, we see that the operator $f \mapsto \sup_{\varepsilon>0} |(\varphi_\varepsilon * f)(x)|$ is of weak type (1,1). Since $\varphi_\varepsilon * f$ converges pointwise (and in fact uniformly) to f when f is in $C_{\text{com}}(\mathbb{R}^N)$, the same argument as for Corollary 6.39 shows that $\lim_{\varepsilon\downarrow 0}(\varphi_\varepsilon * f)(x) = f(x)$ almost everywhere for each f in $L^1(\mathbb{R}^N)$.

EXAMPLE. An example of a function φ as in Corollary 6.42 is $P(x) = \dfrac{1}{1+x^2}$ in $\mathbb{R}^1$. We shall see in Chapter VIII that the function $h(x, y) = (P_y * f)(x)$ for this φ is the natural function on the half plane $y > 0$ in $\mathbb{R}^2$ that is **harmonic**, i.e., has $\dfrac{\partial^2 h}{\partial x^2} + \dfrac{\partial^2 h}{\partial y^2} = 0$, and has boundary value f. Corollary 6.42 says that $h(x, y)$ has $f(x)$ as boundary values almost everywhere if f is in $L^1(\mathbb{R}^1)$.

7. Fourier Series and the Riesz–Fischer Theorem

As mentioned at the beginning of Chapter V, the use of the Riemann integral imposes some limitations on the subject of Fourier series that no longer apply when one uses the Lebesgue integral. In this section we shall redo the elementary theory of Fourier series of Section I.10 with the Lebesgue integral in place, with particular attention to the improved theorems that we obtain. It will be assumed that the reader knows the theory of that section.

The underlying measure space with be $[-\pi, \pi]$ with the σ-algebra of Borel sets and with the measure $\frac{1}{2\pi}\,dx$, where dx is 1-dimensional Lebesgue measure. The complex-valued functions under consideration will be periodic of period 2π, thus assuming the same value at π as at $-\pi$. The spaces L^1, L^2, and L^∞ will refer to this measure space when no other parameters are given. Since the measure of the whole space is finite, these spaces satisfy the inclusions $L^\infty \subseteq L^2 \subseteq L^1$. The functions in L^∞ being essentially bounded, they are certainly integrable and square integrable. The inclusion $L^2 \subseteq L^1$ follows from the Schwarz inequality: $\frac{1}{2\pi}\int_{-\pi}^{\pi} |f|\,1\,dx \le \big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f|^2\,dx\big)^{1/2}\big(\frac{1}{2\pi}\int_{-\pi}^{\pi} 1\,dx\big)^{1/2}$.

There is another way of viewing this measure space that will be especially helpful in relating convolution to the theory. Namely, a periodic function on the line of period 2π may be viewed as a function on the unit circle of $\mathbb{C}$ with the angle as parameter. In fact, convolution is a construction that combines group theory with measure theory when the measure is invariant under the group, and that is why convolution appears more natural on the circle than on $[-\pi, \pi]$. The limits of integration do not have to be written differently from the way they are written on the line, but we must remember that functions are to be extended periodically when we interpret integrands. The factor $\frac{1}{2\pi}$ in front of the measure means that all convolutions of functions are to contain this factor. Thus the definition of **convolution** for nonnegative f and g is

$$(f * g)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)g(y)\,dy.$$

Convolution is commutative and associative on the circle just as in Proposition 6.13, and the various norm estimates of Section 2 are valid in the setting of the circle. The use of dilations has no analog for the circle, and thus the circle has no approximate identities of the form φ_ε.

If f is in L^1, the trigonometric series

$$\sum_{n=-\infty}^{\infty} c_n e^{inx} \qquad \text{with} \qquad c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx$$

is called the **Fourier series** of f. This time we regard the integral as a Lebesgue integral. We write

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx} \qquad \text{and} \qquad s_N(f;x) = \sum_{n=-N}^{N} c_n e^{inx}.$$

A Fourier series can be written also with cosines and sines, and the coefficients a_n and b_n are unchanged from Section I.10.

Theorem 6.43. Let f be in L^2. Among all choices of $d_{-N}, \ldots, d_N$, the expression

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \Big| f(x) - \sum_{n=-N}^{N} d_n e^{inx}\Big|^2\,dx$$

is minimized uniquely by choosing d_n, for all n with $|n| \le N$, to be the Fourier coefficient $c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$. The minimum value is

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx - \sum_{n=-N}^{N} |c_n|^2.$$

PROOF. The proof is the same as for Theorem 1.53.

Corollary 6.44 (Bessel's inequality). If f is in L^2 with $f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$, then

$$\sum_{n=-\infty}^{\infty} |c_n|^2 \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx.$$

In particular, $\sum_{n=-\infty}^{\infty} |c_n|^2$ is finite.

PROOF. The proof is the same as for Corollary 1.54.

Corollary 6.45 (Riemann–Lebesgue Lemma). If f is in L^1 and has Fourier coefficients $\{c_n\}_{n=-\infty}^{\infty}$, then $\lim_{|n|\to\infty} c_n = 0$.

REMARK. Since L^2 is properly contained in L^1, this corollary is not a special case of Corollary 6.44, unlike the situation with the Riemann integral.

PROOF. The result is immediate from Corollary 6.44 in the case of L^2 functions and in particular in the case of continuous functions. Write $c_n(h)$ for the n^{th} Fourier coefficient of any function h. Let $\epsilon > 0$ be given. Choose by Corollary 6.27a a continuous g with $\|f-g\|_1 \le \epsilon/2$. Then choose N such that $|n| \ge N$ implies $|c_n(g)| \le \epsilon/2$. Then $|c_n(f)| \le |c_n(f-g)| + |c_n(g)| \le \epsilon$ since $|c_n(f-g)| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f-g||e^{-inx}|\,dx = \|f-g\|_1 \le \epsilon/2$.

Theorem 6.46 (Dini's test). Let f be in L^1, and fix x in $[-\pi,\pi]$. If there is a constant $\delta > 0$ such that

$$\int_{|t|<\delta} \frac{|f(x+t)-f(x)|}{|t|}\,dt < \infty,$$

then $\lim_N s_n(f;x) = f(x)$.

REMARK. The condition in the corresponding result for the Riemann integral, namely Theorem 1.57, was that $|f(x+t)-f(x)| \le M|t|$ for $|t| < \delta$ and some constant M. The condition in the present theorem is satisfied by $f(x) = \sqrt{|x|}$ at $x = 0$, and the condition in the earlier theorem is not.

PROOF. The proof is the same as for Theorem 1.57 except that we need to appeal to the improved version of the Riemann–Lebesgue Lemma in Corollary 6.45.

Now we work toward a proof of Parseval's Theorem for all of L^2. We need to know about Fourier coefficients of convolutions.

Proposition 6.47. If $f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$ and $g(x) \sim \sum_{n=-\infty}^{\infty} d_n e^{inx}$, then $(f * g)(x) \sim \sum_{n=-\infty}^{\infty} c_n d_n e^{inx}$.

PROOF. This is a consequence of Fubini's Theorem and the translation invariance of Lebesgue measure:

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} (f*g)(x)e^{-inx}\,dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\Big[\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)g(y)e^{-inx}\,dy\Big]\,dx \\
&= \Big(\frac{1}{2\pi}\Big)^2 \int_{-\pi}^{\pi}\Big[\int_{-\pi}^{\pi} f(x-y)g(y)e^{-inx}\,dx\Big]\,dy \\
&= \Big(\frac{1}{2\pi}\Big)^2 \int_{-\pi}^{\pi}\Big[\int_{-\pi}^{\pi} f(x)g(y)e^{-in(x+y)}\,dx\Big]\,dy \\
&= \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx\Big)\Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} g(y)e^{-iny}\,dy\Big).
\end{aligned}$$

The proof of the version of Parseval's Theorem for all of L^2 will make use of the Fejér kernel $K_N(t)$ introduced in Section I.10. We do not need to recall the exact formula for K_N, only the fact that it is a trigonometric polynomial of degree N with the following three properties:

(i) $K_N(x) \geq 0$,
(ii) $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(x)\,dx = 1$,
(iii) for any $\delta > 0$, $\sup_{\delta \leq |x| \leq \pi} K_N(x)$ tends to 0 as n tends to infinity.

These three properties identified K_N as an approximate identity in the setting of periodic functions, and Fejér's Theorem in the form of Theorem 1.59 gave the consequence for convergence at points of continuity of f. With the Lebesgue integral, we get also results about norm convergence in L^1 and L^2.

Theorem 6.48 (Fejér's Theorem). Let f be in L^1. Then

(a) $\lim_N \|K_N * f - f\|_1 = 0$ with no additional hypotheses on f,
(b) $\lim_N \|K_N * f - f\|_2 = 0$ if f is also in L^2,
(c) $\lim_N (K_N * f)(x_0) = f(x_0)$ if f is bounded on $[-\pi, \pi]$ and is continuous at x_0,
(d) the convergence in (c) is uniform for x_0 in E if f is bounded on $[-\pi, \pi]$ and is uniformly continuous at the points of E,
(e) $\lim_N (K_N * f)(x_0) = \frac{1}{2}\big(f(x_0+) + f(x_0-)\big)$ if f is bounded on $[-\pi, \pi]$ and has right and left limits $f(x_0+)$ and $f(x_0-)$ at x_0.

PROOF. For (a) and (b), let $p = 1$ or $p = 2$ as appropriate. Then

$$\begin{aligned}
&\|K_n * f - f\|_p \\
&\qquad = \Big\|\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(t)[f(x-t) - f(x)]\,dt\Big\|_{p,x} && \text{by (ii)} \\
&\qquad \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(t)\|f(x-t) - f(x)\|_{p,x}\,dt && \text{by (i) and Theorem 5.60} \\
&\qquad \leq \sup_{|t|\leq\delta} \|f(x-t) - f(x)\|_{p,x} + 2\,[\sup_{\delta\leq|t|\leq\pi} K_N(t)]\,\|f\|_p.
\end{aligned}$$

Given $\epsilon > 0$, choose δ by Proposition 6.16 to make the first term of the final bound be $< \epsilon/2$, and then choose N_0 by (iii) to make the second term of the final bound be $< \epsilon/2$ for $N \geq N_0$. Then the final bound is $< \epsilon$ for $N \geq N_0$.

Parts (c) and (d) are proved exactly as in Theorem 1.59. For (e), we may assume without loss of generality that $x_0 = 0$ because convolution commutes with translations. If we can prove (e) for a single function g with a jump discontinuity at $x = 0$ equal to the jump for f, then we can apply (c) to $f - g$ and deduce (e) for f. Let us see that such a function g may be taken as a multiple of

$$h(x) = \begin{cases} \frac{1}{2}(\pi - x) & \text{for } 0 < x \leq \pi \\ \frac{1}{2}(-\pi - x) & \text{for } -\pi \leq x < 0. \end{cases}$$

In fact, a computation at the beginning of Section I.10 shows explicitly that the series $\sum_{n=1}^{\infty} (\sin nx)/n$ converges to $h(x)$ for $x \neq 0$, but we do not need this fact. All that we need is that the series $\sum_{n=1}^{\infty} (\sin nx)/n$ is the Fourier series of h, a fact that we can readily check from the definition. The sum of this series at $x = 0$ is manifestly 0, and this sum matches the average of the jumps $\frac{1}{2}\big(\frac{\pi}{2} + \frac{-\pi}{2}\big)$. The Cesàro sums of the series $\sum_{n=1}^{\infty} (\sin nx)/n$ must have the same limit 0, according to Theorem 1.47, and (e) is proved.

Theorem 6.49 (Parseval's Theorem). If f is a function in L^2 with $f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$, then

$$\lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - s_N(f; x)|^2 \, dx = 0$$

and

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx = \sum_{n=-\infty}^{\infty} |c_n|^2.$$

PROOF. From the first conclusion of Theorem 6.43, we obtain $0 \leq \|f - s_N\|_2^2 \leq \|f - (K_N * f)\|_2^2$, and we know from Theorem 6.48b that $\|f - (K_N * f)\|_2^2$ tends to 0. This proves the first formula, and the second formula follows by passing to the limit in the second conclusion of Theorem 6.43.

Corollary 6.50 (uniqueness theorem). If f is in L^1 and has all Fourier coefficients 0, then f is the 0 element in L^1.

PROOF. Proposition 6.47 shows that the Fourier coefficients of $K_N * f$ are $c_n(K_N * F) = c_n(K_N)c_n(f)$, and this is 0 for all n. By Proposition 6.18, $K_N * f$ is continuous, and thus $K_N * f = 0$ by Corollary 1.60. Since $K_N * f$ tends to f in L^1 according to Theorem 6.48a, we conclude that f is the 0 element in L^1.

Now we come to the Riesz–Fischer Theorem, which historically was a great triumph for the Lebesgue integral over the Riemann integral. The result uses the completeness of L^2 and has no counterpart with Riemann integration.

Theorem 6.51 (Riesz–Fischer Theorem). If $\{c_n\}$ is a given doubly infinite sequence of complex numbers with $\sum_{n=-\infty}^{\infty} |c_n|^2 < \infty$, then there exists an f in L^2 whose Fourier series is $\sum_{n=-\infty}^{\infty} c_n e^{inx}$.

PROOF. Define $F_N(x) = \sum_{|n|\le N} c_n e^{inx}$. For $M \ge N$, Parseval's Theorem (Theorem 6.49) gives $\|f_M - f_N\|_2^2 = \sum_{N+1\le|n|\le M} |c_n|^2$, and the right side tends to 0 as M and N tend to infinity because of the convergence of $\sum_{n=-\infty}^{\infty} |c_n|^2$. Thus $\{f_N\}$ is a Cauchy sequence in L^2. By Theorem 5.59, L^2 is complete as a metric space, and thus $\{f_N\}$ converges in L^2. Let f be (a function representing) the limit element in L^2. The inner product in L^2 is a continuous function of the L^2 function in the first variable, and therefore the Fourier coefficients of f satisfy

$$c_n(f) = \big(f, e^{inx}\big) = \lim_N \big(f_N, e^{inx}\big).$$

As soon as N gets to be $\ge |n|$, $\big(f_N, e^{inx}\big)$ equals c_n. Thus $c_n(f) = c_n$ for all n, and f has the required properties.

8. Stieltjes Measures on the Line

A **Stieltjes measure**[2] is a Borel measure on $\mathbb{R}^1$. Lebesgue measure dx is an example, as is any measure $f(x)\,dx$ in which f is nonnegative and Borel measurable and is integrable on every bounded interval. A completely different kind of Stieltjes measure is one that attaches nonnegative weights to countably many points in such a way that the sum of the weights in any bounded interval is finite. In this section we shall see that the Stieltjes measures stand in one-one correspondence with a class of monotone functions on the line that we describe shortly. We shall also obtain an integration-by-parts formula in which a Stieltjes measure plays the role of the derivative of its corresponding monotone function.

If a Stieltjes measure μ is given, we associate to μ the function $F : \mathbb{R}^1 \to \mathbb{R}^1$ defined by

$$F(x) = \begin{cases} -\mu(x, 0] & \text{if } x \le 0, \\ \mu(0, x] & \text{if } x \ge 0. \end{cases}$$

The function F is called the **distribution function** of μ. It has the following properties:[3]

(i) F is nondecreasing, i.e., is **monotone increasing**,

(ii) F is **continuous from the right** in the sense that $F(x_0) = \lim_{x\downarrow x_0} F(x)$ for every x_0 in $\mathbb{R}^1$, i.e., $\lim_n F(x_n) = F(x_0)$ whenever $\{x_n\}_{n\ge1}$ is a sequence tending to x_0 such that $x_n > x_0$ for all $n \ge 1$,

(iii) $F(0) = 0$.

---

[2]Many books, this one included, take Stieltjes measures by definition to occur on the line. However, there is a theory, albeit a somewhat unsatisfactory one, of "Stieltjes measures" in higher-dimensional Euclidean space. It is of interest chiefly in probability theory.

[3]An alternative definition says $F(x)$ equals $-\mu[x, 0)$ and $\mu[0, x)$ in the two cases, and then property (ii) says that F is **continuous from the left**. The choice made here between these alternatives is governed by keeping technicalities to a minimum in Section 10.

Properties (i) and (iii) are immediate from the definition. With (ii), there are two cases according as the limit x_0 is ≤ 0 or > 0, and both cases are settled by the complete additivity of μ.

The measure μ is completely determined by its distribution function F. In fact, the definition of F forces $\mu((a,b]) = F(b) - F(a)$, and Proposition 6.6 implies that μ is determined as a Borel measure by this formula.

Theorem 6.52. The Stieltjes measures μ stand in one-one correspondence with the functions $F : \mathbb{R}^1 \to \mathbb{R}^1$ satisfying (i), (ii), and (iii), the correspondence being that F is the distribution function of μ.

PROOF. We have seen that each μ leads to an F and that F uniquely determines μ. We need to see that every F satisfying (i), (ii), and (iii) arises from some μ. If such a function F is given, we define a set function μ on bounded intervals by

$$\begin{aligned}
\mu((a,b]) &= F(b) - F(a), \\
\mu((a,b)) &= \lim_n F(b - \tfrac{1}{n}) - F(a), \\
\mu([a,b]) &= F(b) - \lim_n F(a - \tfrac{1}{n}), \\
\mu([a,b)) &= \lim_n F(b - \tfrac{1}{n}) - \lim_n F(a - \tfrac{1}{n}).
\end{aligned}$$

We extend μ to the ring $\mathcal{R}$ of elementary subsets of $\mathbb{R}^1$, i.e., the ring of all finite disjoint unions of bounded intervals, by setting μ of a finite disjoint union of bounded intervals equal to the sum of the values of μ on each of the intervals, just as with Lebesgue measure in Example 4 at the end of Section V.1.

To see that μ is unambiguously defined and is additive on $\mathcal{R}$, we readily reduce matters, just as with Lebesgue measure, to showing that if an interval is decomposed into the union of two smaller intervals, then μ of the union is the sum of μ of the components. Thus let $a \le b \le c$, and let an interval I from a to c be the union of an interval from a to b and an interval from b to c. If the interval I from a to c is (a,c), then the two possible cases are handled by

$$\begin{aligned}
\mu((a,b)) + \mu([b,c)) &= \lim_n F(b - \tfrac{1}{n}) - F(a) + \lim_n F(c - \tfrac{1}{n}) - \lim_n F(b - \tfrac{1}{n}) \\
&= \mu((a,c))
\end{aligned}$$

and

$$\mu((a,b]) + \mu((b,c)) = F(b) - F(a) + \lim_n F(c - \tfrac{1}{n}) - F(b) = \mu((a,c)).$$

If the interval I from a to c has one or both endpoints present, then the computation is the same except that $F(a)$ is replaced by $\lim_n F(a - \frac{1}{n})$ if a is in I and

$\lim_n F(c - \frac{1}{n})$ is replaced by $F(c)$ if c is in I. Thus μ is unambiguously defined, and it is nonnegative and additive.

The next step is to prove, just as with Lebesgue measure in Section V.1, that μ is regular on $\mathcal{R}$ in the sense that for each E in $\mathcal{R}$ and $\epsilon > 0$, we can find a compact K in $\mathcal{R}$ and an open U in $\mathcal{R}$ such that $K \subseteq E \subseteq U$, $m(K) \geq m(E) - \epsilon$, and $m(U) \leq m(E) + \epsilon$. As with Lebesgue measure, the proof comes down to the case that E is a single interval, and this time there are four subcases. Choose n large enough so that $\frac{2}{n} < \epsilon$, and then

$$\begin{aligned}
&\text{for } [a, b), \quad \text{take } K = [a, b - \tfrac{1}{n}] \text{ and } U = (a - \tfrac{1}{n}, b),\\
&\text{for } [a, b], \quad \text{take } K = [a, b] \text{ and } U = (a - \tfrac{1}{n}, b + \tfrac{1}{n}),\\
&\text{for } (a, b], \quad \text{take } K = [a + \tfrac{1}{n}, b] \text{ and } U = (a, b + \tfrac{1}{n}),\\
&\text{for } (a, b), \quad \text{take } K = [a + \tfrac{1}{n}, b - \tfrac{1}{n}] \text{ and } U = (a, b).
\end{aligned}$$

An exception occurs in the definition of K if the listed left endpoint of K exceeds the listed right endpoint of K, and then K is defined to be empty. Each of these definitions contains a parameter n; if we write K_n and U_n for the corresponding sets K and U, then we can check from the definitions and property (ii) of the function F that $\lim_n \mu(K_n) = \mu(E)$ and $\lim_n \mu(U_n) = \mu(E)$. The regularity condition for E follows from these limit relations.

The next step is to prove that μ is completely additive on $\mathcal{R}$ by imitating the proof for Lebesgue measure. In fact, the proof of Proposition 5.4 applies word-for-word except that m has to be changed to μ throughout and the word "proposition" in the last sentence of the proof should be changed to "complete additivity."

Then μ extends to a measure on the Borel sets by Theorem 5.5. The extended measure is σ-finite on $\mathbb{R}^1$ because $\mathbb{R}^1$ is the countable union of bounded intervals and μ is finite on every bounded interval.

Finally we need to show that the distribution function G of μ is equal to F. Our definitions make

$$G(x) = \begin{cases} -\mu((x, 0]) = -(F(0) - F(x)) = F(x) & \text{if } x \leq 0,\\ \mu((0, x]) = F(x) - F(0) = F(x) & \text{if } x \geq 0. \end{cases}$$

Thus $G = F$.

EXAMPLES.

(1) Let F be any continuous distribution function that has a continuous derivative f except possibly at finitely many points. If x is a point of continuity of f, then the Fundamental Theorem of Calculus (Theorem 1.32) gives

$\frac{d}{dx}\int_0^x f(t)\,dt = f(x)$. Put

$$G(x) = \begin{cases} -\int_x^0 f(t)\,dt & \text{if } x \le 0, \\ \int_0^x f(t)\,dt & \text{if } x \ge 0. \end{cases}$$

Then G is a continuous distribution function, and the formula for the derivative of the integral shows that $F'(x) = G'(x)$ except at finitely many points. Recursive application of the Mean Value Theorem, starting from $x = 0$, to $F - G$ on intervals having $F' - G' = 0$ in their interiors, shows that $F = G$ everywhere. The Stieltjes measure μ associated to F, by the uniqueness in Theorem 6.52, is given by

$$\mu(E) = \int_E f(t)\,dt.$$

The special case with $F(x)$ identically equal to x has f identically equal to 1, and the measure is just Lebesgue measure.

(2) The function F with

$$F(x) = \begin{cases} 0 & \text{for } x \ge 0, \\ -1 & \text{for } x < 0, \end{cases}$$

has the three properties of a distribution function, and the associated measure μ is a point mass assigning weight 1 to $x = 0$. The measure μ takes the value 1 on every Borel set containing 0 and takes the value 0 on every Borel set not containing 0. This measure is sometimes called the **delta measure** at 0 or "delta mass" at 0. Whenever a Stieltjes measure ν has $\nu(\{p\}) > 0$ for some p in $\mathbb{R}^1$, we say that ν contains a **point mass** at p of weight $\nu(\{p\})$. Then ν is the sum of a point mass at p of weight $\nu(\{p\})$ and a Stieltjes measure containing no point mass at p.

(3) Let $\{x_n\}$ be a sequence in $\mathbb{R}$. For example, $\{x_n\}$ could be an enumeration of the rationals. Let $\{w_n\}$ be a sequence of positive numbers with $\sum w_n < \infty$, and define

$$F(x) = \begin{cases} \displaystyle\sum_{\{n \mid 0 < x_n \le x\}} w_n & \text{for } x \ge 0, \\ \displaystyle -\sum_{\{n \mid x < x_n \le 0\}} w_n & \text{for } x < 0. \end{cases}$$

Then F satisfies (i), (ii), and (iii), and hence F is the distribution function of some Stieltjes measure μ. The measure is given by

$$\mu((a, b]) = \sum_{\{n \mid a < x_n \le b\}} w_n.$$

It is a countable sum of point masses. The function F, though monotone increasing, is discontinuous at every x_n, and this set is allowed to be dense.

(4) This example will be a nonzero Stieltjes measure that is carried on a Borel set of Lebesgue measure 0 and yet has a continuous distribution function. We start from the standard Cantor set C in $[0, 1]$ described in detail in Section II.9. This set is compact and is obtained as the intersection of a sequence $\{C_n\}$ of sets with each C_n consisting of the finite union of closed bounded intervals. The set C_0 is $[0, 1]$, and C_{n+1} is obtained from C_n by removing the open middle third of each of the

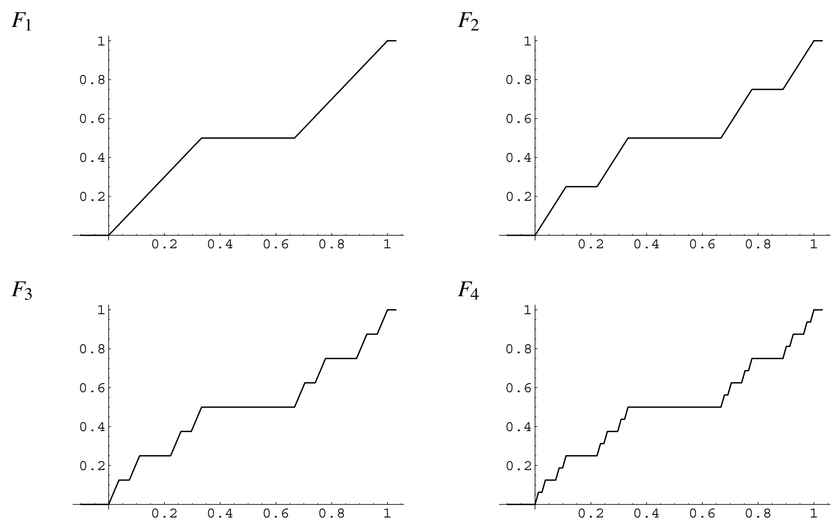


FIGURE 6.1. Construction of a Cantor function F. Graphs of approximations F_1, F_2, F_3, F_4 to F.

constituent closed intervals of C_n. The Lebesgue measure of C_n is $(2/3)^n$, and thus C has Lebesgue measure $\lim_n (2/3)^n = 0$. The measure μ we construct will have $\mu(C) = 1$ and $\mu(C^c) = 0$, yet it will assign 0 measure to every one-point set. The properties that are needed of the corresponding distribution function F so that μ has these properties are that F is continuous, F is 0 for $x \leq 0$, F is 1 for $x \geq 1$, and F is constant on every open interval I of $[0, 1] - C$, i.e., on every open interval of every $[0, 1] - C_n$. This condition will make $\mu(I) = 0$ for all such I. Since the metric space $[0, 1] - C$ has a countable base, it is the union of countably many such open intervals I, and thus $\mu([0, 1] - C) = 0$. Since F is constant for $x \leq 0$ and for $x \geq 1$, μ is 0 on $(-\infty, 0)$ and $(1, +\infty)$ as well, and thus $\mu(C^c) = 0$. To obtain the distribution function F, we construct a sequence of approximating functions F_n and show that the sequence is uniformly Cauchy. The set $C_n^c \cap [0, 1]$ is the union of $2^n - 1$ disjoint open intervals. On the k^{th} such interval we define F_n to be $k2^{-n}$. We let $F_n(x) = 0$ for $x \leq 0$ and $F(x) = 1$

for $x \geq 1$. On the complementary closed intervals, define F_n in any fashion that makes F_n monotone increasing and continuous. Graphs of F_1 through F_4 are shown in Figure 6.1 with the interpolation in the graphs done by straight lines. The result is that

$$|F_n(x) - F_{n+k}(x)| \leq 2^{-n}$$

for all x. Hence $\{F_n\}$ is uniformly Cauchy and therefore uniformly convergent. Let F be the limit function. The function F continuous by Theorem 1.21, and it is monotone increasing, satisfies $F(0) = 0$, and is constant on every open interval contained in C^c. According to Problem 15 at the end of the chapter, it is independent of the method of interpolation used in constructing the F_n's. The function F is called the **Cantor function** corresponding to the standard Cantor set.

The most general monotone increasing function F on $\mathbb{R}^1$ is not far from being the distribution function of some Stieltjes measure. In the first place the monotonicity of F implies that F has left and right limits at every point, and consequently its only discontinuities are jumps. There can be only countably many such jumps: in fact, if there were uncountably many jump discontinuities, there would be uncountably many in some bounded interval, and that interval would contain uncountably many of magnitude at least $1/n$ for some integer n; hence F would have to be unbounded on that bounded interval. Let us define a function F_1 by $F_1(x) = \lim_{t \downarrow x} F(t)$. This is well defined, since F has right limits at every point, and we have $F(x) = F_1(x)$ except on a countable set. If we define $F_2(x) = F_1(x) - F_1(0)$, then F_2 satisfies the three defining properties (i), (ii), (iii) of a distribution function. If μ is the Stieltjes measure corresponding to F_2 under Theorem 6.52, we call μ the **associated Stieltjes measure** for F.

Theorem 6.53 (integration by parts). Let $a < b$, let F be a monotone increasing function on $\mathbb{R}^1$ that is continuous from the right at a and b, and let μ be the associated Stieltjes measure. If G is a C^1 complex-valued function on $[a, b]$ with derivative g, then

$$\int_a^b F(x)g(x)\,dx = G(b)F(b) - G(a)F(a) - \int_{(a,b]} G\,d\mu.$$

PROOF. Without loss of generality, we may assume that G is real-valued. Let F_2 be the distribution function of μ. By construction of F_2, there is a constant c such that $F - F_2 = c$ except possibly at points of discontinuity of F, and the set S of such points S within $[a, b]$ is countable. This exceptional countable set S does not contain a or b, since F and F_2 are continuous from the right at a and b.

We have

$$\begin{aligned}\int_a^b (F - F_2)g\,dx &= \int_S (F - F_2)g\,dx + \int_a^b cg(x)\,dx \\ &= c\int_a^b g(x)\,dx = c(G(b) - G(a))\end{aligned}$$

and also

$$G(b)(F(b)-F_2(b))-G(a)(F(a)-F_2(a)) = G(b)c-G(a)c = c(G(b)-G(a)).$$

Thus

$$\int_a^b (F - F_2)g\,dx = G(b)(F(b) - F_2(b)) - G(a)(F(a) - F_2(a)).$$

Comparing this formula with the formula in the statement of the theorem, we see that if the theorem holds for F_2, then it holds for F. Changing notation, we may therefore assume that F is the distribution function of μ.

Let P be a partition $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ of $[a, b]$ with mesh to be specified. For $1 \le i \le n$, we use the Mean Value Theorem to choose $t_i \in (x_{i-1}, x_i)$ with $G(x_i) - G(x_{i-1}) = g(t_i)(x_i - x_{i-1})$. We can do so since we have assumed that G is real valued. Then we have

$$F(x_i)g(t_i)(x_i - x_{i-1}) = F(x_i)G(x_i) - F(x_i)G(x_{i-1})$$

and

$$\begin{aligned}\sum_{i=1}^n & F(x_i)g(t_i)(x_i - x_{i-1}) \\ &= F(x_n)G(x_n) + \sum_{i=2}^n F(x_{i-1})G(x_{i-1}) - \sum_{i=1}^n F(x_i)G(x_{i-1}) \\ &= F(x_n)G(x_n) - F(x_0)G(x_0) - \sum_{i=1}^n G(x_{i-1})(F(x_i) - F(x_{i-1})) \\ &= F(b)G(b) - F(a)G(a) - \sum_{i=1}^n G(x_{i-1})(F(x_i) - F(x_{i-1})).\end{aligned}$$

We shall show for small enough mesh that $\sum_{i=1}^n F(x_i)g(t_i)(x_i - x_{i-1})$ is close to $\int_a^b F(x)g(x)\,dx$ and that $\sum_{i=1}^n G(x_{i-1})(F(x_i) - F(x_{i-1}))$ is close to $\int_{(a,b]} G\,d\mu$.

Let M be an upper bound for $|g|$ and $|F|$ on $[a, b]$, and let $\epsilon > 0$ be given. Choose a number $\delta > 0$ by uniform continuity of G and g such that $|x - x'| < \delta$ implies $|G(x) - G(x')| < \epsilon/M$ and $|g(x) - g(x')| < \epsilon/(M(b - a))$, as well as another condition to be specified. If the mesh of the partition is $< \delta$, then

$$
\begin{aligned}
\Big|\sum_{i=1}^{n} G(x_{i-1})(F(x_i) - F(x_{i-1})) - \textstyle\int_{(a,b]} G\,d\mu\Big| \\
&= \Big|\sum_{i=1}^{n} \textstyle\int_{(x_{i-1},x_i]}[G(x_{i-1}) - G(x)]\,d\mu\Big| \\
&\le \sum_{i=1}^{n} \textstyle\int_{(x_{i-1},x_i]} |G(x_{i-1}) - G(x)|\,d\mu \\
&\le \sum_{i=1}^{n} \textstyle\int_{(x_{i-1},x_i]}(\epsilon/M)\,d\mu \\
&= (\epsilon/M)(F(b) - F(a)) \\
&\le 2\epsilon \qquad\qquad \text{since } |F| \le M.
\end{aligned}
$$

Also,

$$
\begin{aligned}
&\Big|\sum_{i=1}^{n} F(x_i)g(t_i)(x_i - x_{i-1}) - \textstyle\int_a^b F(x)g(x)\,dx\Big| \\
&\quad= \Big|\sum_{i=1}^{n} \textstyle\int_{(x_{i-1},x_i]} \big(F(x_i)(g(t_i) - g(x)) + (F(x_i) - F(x))g(x)\big)\,dx\Big| \\
&\quad\le \sum_{i=1}^{n} \textstyle\int_{(x_{i-1},x_i]} |F(x_i)|\,|g(t_i) - g(x)|\,dx + \sum_{i=1}^{n} \textstyle\int_{(x_{i-1},x_i]} |F(x_i) - F(x)|\,|g(x)|\,dx \\
&\quad\le \sum_{i=1}^{n} \textstyle\int_{(x_{i-1},x_i]} M(\epsilon/(M(b-a)))\,dx + \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})|M\delta \\
&\quad= \epsilon + (F(b) - F(a))M\delta \qquad \text{by monotonicity of } F \\
&\quad\le \epsilon + 2M^2\delta.
\end{aligned}
$$

Thus if δ satisfies the additional condition that $\delta < \epsilon/(2M^2)$, then the absolute value of the difference of the two sides in the formula of the theorem is $< 2\epsilon + 2\epsilon = 4\epsilon$. This completes the proof.

9. Fourier Series and the Dirichlet–Jordan Theorem

A real-valued function f on a bounded interval $[a, b]$ is said to be of **bounded variation** on $[a, b]$ if there is a constant M such that every partition

$$P: \quad a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

has

$$\sup_P \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \le M.$$

Let us write $\|f\|_{BV}$ for the least M such that this inequality holds. The set of functions f of bounded variation on $[a, b]$ is a pseudo normed linear space in the sense of Section V.9, the pseudonorm being $\|\cdot\|_{BV}$. The only functions f with $\|f\|_{BV} = 0$ are the constants.

Examples of functions of bounded variation are furnished by arbitrary bounded monotone functions and by any function with a continuous derivative. In fact, if f is monotone increasing, then f is of bounded variation with $\|f\|_{BV} = f(b) - f(a)$. If f has f' continuous, then the Mean Value Theorem gives

$$\begin{aligned}\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| &= \sum_{i=1}^{n} |f'(t_i)|(x_i - x_{i-1}) \qquad \text{with } x_{i-1} < t_i < x_i \\ &\le \|f'\|_{\sup} \sum_{i=1}^{n} (x_i - x_{i-1}) \\ &= \|f'\|_{\sup}(b-a),\end{aligned}$$

and we see that f is of bounded variation with $\|f\|_{BV} \le \|f'\|_{\sup}(b-a)$.

Let us associate two functions on $[a, b]$ to f if f is of bounded variation. For a real number r, define $r^+ = \max\{r, 0\}$ and $r^- = -\min\{r, 0\}$, so that $r = r^+ - r^-$ and $|r| = r^+ + r^-$. The functions are the **positive** and **negative variations** of f, given by

$$V^+(f)(x) = \sup_{\substack{P \text{ with } x_0=a \\ \text{and } x_n=x}} \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)^+,$$

$$V^-(f)(x) = \sup_{\substack{P \text{ with } x_0=a \\ \text{and } x_n=x}} \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)^-,$$

the supremum in each case being taken over all partitions of $[a, x]$.

Proposition 6.54. If f is of bounded variation on $[a, b]$, then $V^+(f)$ and $V^-(f)$ are monotone increasing functions such that

$$f(x) = f(a) + V^+(f)(x) - V^-(f)(x)$$

for all x in $[a, b]$. In particular, f is the difference of two monotone increasing functions.

REMARK. Since monotone functions have left and right limits at each point, it follows that every function f of bounded variation has left and right limits at each point. We denote these by $f(x-)$ and $f(x+)$, respectively. The function f is continuous from the left at x if and only if $f(x-) = f(x)$, and it is continuous from the right at x if and only if $f(x) = f(x+)$.

PROOF. It is evident from the definitions that $V^+(f)$ and $V^-(f)$ are monotone increasing. Fix x, and let P be a partition of $[a, x]$. Then

$$\begin{aligned} f(x) - f(a) &= \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big) \\ &= \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)^+ - \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)^-. \end{aligned}$$

Hence

$$\begin{aligned} \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)^+ &= \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)^- + \big(f(x) - f(a)\big) \\ &\le V^-(f)(x) + \big(f(x) - f(a)\big), \end{aligned}$$

and

$$V^+(f)(x) \le V^-(f)(x) + \big(f(x) - f(a)\big).$$

Also,

$$\begin{aligned} \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)^- &= \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big)^+ - \big(f(x) - f(a)\big) \\ &\le V^+(f)(x) - \big(f(x) - f(a)\big), \end{aligned}$$

and

$$V^-(f)(x) \le V^+(f)(x) - \big(f(x) - f(a)\big).$$

Therefore

$$f(x) - f(a) = V^+(f)(x) - V^-(f)(x),$$

and the proof is complete.

Theorem 6.55 (Dirichlet–Jordan Theorem). If f is a function of bounded variation on $[-\pi, \pi]$, then the Fourier series of f converges at each point to $\frac{1}{2}\big(f(x-) + f(x+)\big)$ and it converges uniformly to $f(x)$ on any compact set on which the periodic extension of f is continuous.

By way of preparation, it will be convenient to extend the definition of Fourier series to allow integrable functions to be replaced by more general Borel measures. If μ is a Borel measure on $[-\pi, \pi]$, we want to be able to regard μ as periodic. One way to proceed would be to insist that μ really be a measure on the circle group, hence be defined on $(-\pi, \pi]$. Alternatively, we could insist that any point mass contributing to μ at $-\pi$ be matched by an equal point mass for μ at π. A way of avoiding point masses contributing at the endpoints is to change the interval $[-\pi, \pi]$ to a suitable $[c-\pi, c+\pi]$; we can find a number c with no point masses at the ends because only countably many point masses can contribute to μ and still have μ be a finite measure. In any event, we define the **Fourier–Stieltjes series** of μ to be the series

$$\sum_{n=-\infty}^{\infty} c_n e^{inx} \qquad \text{with} \qquad c_n = \int_{(-\pi,\pi]} e^{-inx}\, d\mu(x).$$

The usual factor of $\frac{1}{2\pi}$ is dropped because we identify an integrable function $f \geq 0$ with the measure $\frac{1}{2\pi} f\, dx$ when making the generalization. From the definition of the Fourier–Stieltjes coefficients, we see immediately that $|c_n| \leq \mu((-\pi, \pi])$; hence the coefficients are bounded.

PROOF OF THEOREM 6.55. We take the given function f to be periodic of period 2π. On some closed interval $[a, b]$ containing $[-\pi, \pi]$ in its interior, let us decompose f according to Proposition 6.54 as $f = f(a) + V^+(f) - V^-(f)$. It is then enough to prove the theorem for the monotone increasing functions $f(a) + V^+(f)$ and $V^-(f)$ separately. These functions need to be extended to all of $\mathbb{R}^1$, and we may make that extension by taking them to be constant to the left of $[a, b]$ and to the right of $[a, b]$.

Changing notation in the theorem, we may assume from the outset that f is monotone and bounded, though no longer periodic. Neither the Fourier coefficients of f nor the hoped-for values of the sum of the Fourier series are changed if we adjust f on a subset of the countable set where f is discontinuous. Thus we may assume without loss of generality that f is continuous from the right at every point. Let $f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$ be its Fourier series.

Let μ be the Stieltjes measure associated to f. Applying integration by parts (Theorem 6.53) on the interval $[-\pi, \pi]$ with $G(x) = e^{-inx}$ and $g(x) = -ine^{-inx}$, we obtain

$$\int_{-\pi}^{\pi} f(x)(-in)e^{-inx}\, dx = e^{-in\pi} f(\pi) - e^{-in(-\pi)} f(-\pi) - \int_{(-\pi,\pi]} e^{-inx}\, d\mu(x).$$

The left side is $-2\pi i n c_n$, and the right side is the sum of two bounded terms and the negative of a Fourier–Stieltjes coefficient of μ. These Fourier–Stieltjes coefficients are bounded, and hence $|c_n| \leq C/|n|$ for some constant C.

Let $s_N(x) = \sum_{n=-N}^{N} c_n e^{inx}$ be the N^{th} partial sum of the Fourier series of f, and let $\sigma_N(x) = \frac{1}{N+1} \sum_{k=0}^{N} s_k(x)$ be the N^{th} Cesàro sum. We know that $\sigma_N(x) = (K_N * f)(x)$, where K_N is the Fejér kernel. Fejér's Theorem (Theorem 6.48) shows that $\lim_N \sigma_N(x) = \frac{1}{2}\big(f(x-) + f(x+)\big)$ for all x and that $\lim_N \sigma_N(x) = f(x)$ uniformly on any compact set of points where f is continuous. The Tauberian theorem stated as Proposition 1.50 allows us to conclude that $s_N(x)$ converges and has the same limit as $\sigma_N(x)$ if it is shown that the sequence $n(c_n e^{inx} + c_{-n}e^{-inx})$ is bounded for $n > 0$. But this boundedness is immediate from the estimate $|c_n| \le C/|n|$ for the Fourier coefficients of f.

10. Distribution Functions

This section concerns the computation of integrals. A measure space $(X, \mathcal{A}, \rho)$ will be fixed throughout. A need to estimate integrals arises in two quite distinct situations, and the emphasis is different for the two situations. One is in connection with problems in Fourier analysis and differential equations, and the underlying measure space typically has X equal to $\mathbb{R}^N$, $\mathcal{A}$ equal to the σ-algebra of Borel sets, and ρ equal to Lebesgue measure. The other is in connection with probability theory, and the underlying measure space is typically a complicated space with $\rho(X) = 1$. Although the word "distribution" acquires multiple meanings in the process, the theory can begin at the same point in the two cases.

Let $f : X \to \mathbb{R}$ be a measurable function. We define a measure μ_f on the Borel sets of $\mathbb{R}$ and a function $\lambda_f : (0, +\infty) \to [0, +\infty]$ by

$$\begin{aligned}\mu_f(E) &= \rho(f^{-1}(E)) = \rho\big(\{x \in X \bigm| f(x) \in E\}\big) \quad \text{for each Borel set } E,\\ \lambda_f(\xi) &= \rho(|f|^{-1}((\xi, +\infty))) = \rho\big(\{x \in X \bigm| |f(x)| > \xi\}\big).\end{aligned}$$

Proposition 6.56. If $f : X \to \mathbb{R}$ is a measurable function, then

(a) $\int_X \Phi(f(x))\, d\rho(x) = \int_{\mathbb{R}} \Phi(t)\, d\mu_f(t)$ for every nonnegative Borel measurable function $\Phi : \mathbb{R} \to \mathbb{R}$,

(b) $\int_X \Phi(|f(x)|)\, d\rho(x) = \int_0^\infty \lambda_f(\xi)\varphi(\xi)\, d\xi$ whenever $\varphi(\xi)\, d\xi$ is a Stieltjes measure on $\mathbb{R}^1$ and Φ is its distribution function.

PROOF. In (a), when Φ is an indicator function I_E, the two sides of the identity are $\rho(f^{-1}(E))$ and $\mu_f(E)$, and these are equal by definition of μ_f. We can pass to nonnegative simple functions by linearity and then to general nonnegative Borel measurable functions Φ by monotone convergence.

In (b), when f is a nonnegative simple function s, let $s = \sum_{k=1}^n c_k I_{E_k}$ be the canonical expansion of s as a linear combination of indicator functions, with the c_j's arranged so that $c_1 > c_2 > \cdots > c_n \ge 0$. Put $c_{n+1} = 0$. Then we have

$$\begin{aligned}\int_0^\infty \lambda_f(\xi)\varphi(\xi)\, d\xi &= \sum_{k=1}^n \int_{c_{k+1}}^{c_k} \rho\big(\textstyle\bigcup_{j=1}^k E_j\big)\varphi(\xi)\, d\xi\\ &= \sum_{k=1}^n \rho\big(\textstyle\bigcup_{j=1}^k E_j\big)[\Phi(c_k) - \Phi(c_{k+1})]\\ &= \sum_{k=1}^n \sum_{j=1}^k \rho(E_j)[\Phi(c_k) - \Phi(c_{k+1})]\end{aligned}$$

$$= \sum_{j=1}^n \sum_{k=j}^n \rho(E_j)[\Phi(c_k) - \Phi(c_{k+1})]$$
$$= \sum_{j=1}^n \rho(E_j)\Phi(c_j) \qquad \text{since } \Phi(0) = 0$$
$$= \sum_{j=1}^n \int_{E_j} \Phi(s(x))\, d\rho(x)$$
$$= \int_X \Phi(s(x))\, d\rho(x).$$

This proves (b) for nonnegative simple functions f. For a general measurable $|f|$ on X, choose an increasing sequence of nonnegative simple functions s_n with pointwise limit $|f|$. The definition of Φ in terms of φ makes Φ monotone increasing and continuous, and thus $\Phi(s_n(x))$ increases to $\Phi(|f(x)|)$. Also, the set $\{x \in X \mid |f(x)| > \xi\}$, for each fixed ξ, is the increasing union of the sets $\{x \in X \mid s_n(x) > \xi\}$, and thus $\lambda_{s_n}(\xi) = \rho(\{x \in X \mid s_n(x) > \xi\})$ increases to $\lambda_f(\xi) = \rho(\{x \in X \mid |f(x)| > \xi\})$ for each ξ. Hence we can pass to the limit in the identity for each s_n and obtain the identity for $|f|$ by monotone convergence. This proves (b) for a general measurable $|f|$.

For applications to Fourier analysis and differential equations, it is (b) that is important, and the function Φ of most interest is $\Phi(t) = t^p$ with $0 < p < +\infty$. The formula in this case is

$$\boxed{\int_X |f(x)|^p\, d\rho(x) = p \int_0^\infty \lambda_f(\xi)\xi^{p-1}\, d\xi.}$$

Somewhat unfortunately, the function λ_f is called the **distribution function** of f; the term does not conflict with the notion of the "distribution function" of a Stieltjes measure as long as one does not make any associations between functions and measures.

A special case of the displayed formula is that X is $\mathbb{R}^N$, ρ is Lebesgue measure, and p is 1. In this case the formula simplifies to $\int_{\mathbb{R}^N} |f(x)|\, dx = \int_0^\infty \lambda_f(\xi)\, d\xi$, a formula that was mentioned after the statement of the Hardy–Littlewood Maximal Theorem (Theorem 6.38).

The displayed formula shows that $\int_X |f|^p\, d\rho$ can be computed from the function λ_f, and it is apparent that the integral cannot be finite if $\lambda_f(\xi)$ is everywhere $\geq$ some positive multiple of ξ^{-p}. This observation can be improved upon without the aid of Proposition 6.56 in the following way. We have $\int_X |f(x)|^p\, d\rho(x) \geq \int_{\{x\in X \mid |f(x)|>\xi\}} |f(x)|^p\, d\rho(x) \geq \xi^p \rho(\{x \in X \mid |f(x)| > \xi\}$. Thus we obtain

$$\boxed{\rho(\{x \in X \mid |f(x)| > \xi\} \leq \frac{\int_X |f|^p\, d\rho}{\xi^p},}$$

an inequality that goes under the name **Chebyshev's inequality**.

11. Problems

1. Let S^1 be the unit circle of $\mathbb{C}$, let T be the subgroup of elements of finite order, and let E be a subset of S^1 that contains exactly one element of each coset in S^1/T. (Such a set E exists by the Axiom of Choice.) Prove that E is not a Lebesgue measurable subset of the circle and therefore that the corresponding subset of $(-\pi, \pi]$ is not Lebesgue measurable on $\mathbb{R}^1$.

2. Let φ be the mapping given explicitly in Section 5 that allows one to substitute in an expression in Cartesian coordinates and obtain an expression in spherical coordinates. Let U be the domain of φ. Prove that
   (a) the determinant factor in the change-of-variables formula is given by
   $$|\det \varphi'| = r^{N-1} \sin^{N-2}\theta_1 \sin^{N-3}\theta_2 \cdots \sin\theta_{N-2},$$
   (b) φ is one-one on U,
   (c) the complement of $\varphi(U)$ in $\mathbb{R}^N$ is a lower-dimensional set.

3. Let L be a nonsingular N-by-N real matrix. Prove that
   $$\int_{\mathbb{R}^N} f(Lx)\,dx = |\det L|^{-1} \int_{\mathbb{R}^N} f(x)\,dx$$
   for every nonnegative Borel measurable function f.

4. Let M_N denote the N^2-dimensional Euclidean space of all real N-by-N matrices, and let dx refer to its Lebesgue measure. Prove that
   $$\int_{M_N} f(yx)\,\frac{dx}{|\det x|^N} = \int_{M_N} f(x)\,\frac{dx}{|\det x|^N}$$
   for each nonsingular matrix y and Borel measurable function $f \geq 0$. In the formula, yx is the matrix product of y and x.

5. Fix α with $0 < \alpha < 1$. Suppose $f : \mathbb{R} \to \mathbb{C}$ is periodic of period 2π, is smooth except at multiples of 2π, and satisfies the inequalities $|f(x)| \leq C|x|^\alpha$, $|f'(x)| \leq C|x|^{\alpha-1}$, and $|f''(x)| \leq C|x|^{\alpha-2}$ for $|x| \leq 1$.
   (a) By breaking the integral at $|x| = 1/|n|$, prove that the Fourier coefficients c_n of f satisfy $|c_n| \leq K/|n|^{1+\alpha}$.
   (b) How can one conclude from (a) that the Fourier series of f converges uniformly? Why is the limit equal to f?
   (c) Prove or disprove: The real and imaginary parts of the function f are of bounded variation on every bounded interval.

6. Let μ be a nonzero measure on the σ-algebra of all subsets of $\mathbb{R}^1$ assigning to each set either measure 0 or measure 1. Prove that μ is a point mass.

7. Determine all Stieltjes measures $\nu \neq 0$ on the line with
$$\int_{\mathbb{R}^1} fg\,d\nu = \Big(\int_{\mathbb{R}_1} f\,d\nu\Big)\Big(\int_{\mathbb{R}^1} g\,d\nu\Big)$$
for all continuous nonnegative functions f and g.

Problems 8–10 make use of Fubini's Theorem in unexpected ways.

8. (a) Show that the complement of any Lebesgue measurable set of Lebesgue measure 0 in $\mathbb{R}^N$ is dense.
   (b) Let μ be a Stieltjes measure on the line, and let E be a Borel set in $\mathbb{R}^1$ with Lebesgue measure 0. Prove that $\mu(E+t) = 0$ for almost every t with respect to Lebesgue measure.
   (c) Suppose that a Stieltjes measure μ on the line satisfies $\lim_{t\to 0} \mu(E+t) = \mu(E)$ for each bounded Borel set E in $\mathbb{R}^1$. Prove that $\mu(E) = 0$ for every Borel set E of Lebesgue measure 0.

9. In potential theory a positive charge on $\mathbb{R}^3$ is by definition any finite Borel measure μ, and its potential h is the function $h(x) = \int_{\mathbb{R}^3} \frac{d\mu(y)}{|x-y|}$. Prove that the potential is finite almost everywhere with respect to Lebesgue measure.

10. Let $P(x_1, \ldots, x_n)$ be a real-valued polynomial on $\mathbb{R}^n$ that is not identically 0. Prove by induction that the set in $\mathbb{R}^n$ where $P = 0$ has Lebesgue measure 0.

Problems 11–14 concern the gamma function and some associated changes of variables.

11. Prove that
$$\int_0^1 t^{x-1}(1-t)^{y-1}\,dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$
by starting from the product of $\Gamma(x+y)$ and the left side, substituting for $\Gamma(x+y)$, making a change of variables, using Fubini's Theorem, and making another change of variables.

12. By evaluating the integral $\int_{\mathbb{R}^N} e^{-\pi|x|^2}\,dx$ first in Cartesian coordinates by means of Proposition 6.33 and then in spherical coordinates by means of the change-of-variables formula for multiple integrals, obtain an expression for the area $\Omega_{N-1} = \int_{S^{N-1}} d\omega$ of the sphere S^{N-1}. Express the answer in terms of a value of the gamma function.

13. Let I be the "cube" of all $u = (u_1, \ldots, u_n)$ in $\mathbb{R}^n$ with $0 < u_i < 1$ for all i, and let S be the "simplex" of all $x = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ with $x_i > 0$ for all i and $\sum_{i=1}^n x_i < 1$. Define $x = \varphi(u)$ by

$$\begin{aligned} x_1 &= u_1, \\ x_2 &= (1-u_1)u_2, \\ &\vdots \\ x_n &= (1-u_1)\cdots(1-u_{n-1})u_n. \end{aligned}$$

(a) Prove that $\sum_{i=1}^n x_i = 1 - \prod_{i=1}^n (1-u_i)$.
(b) Prove that φ maps I one-one onto S, with inverse given by

$$u_i = \frac{x_i}{1 - x_1 - \cdots - x_{i-1}}.$$

(c) Prove that $|\det \varphi'(u)| = (1-u_1)^{n-1}(1-u_2)^{n-2}\cdots(1-u_{n-1})$ and that

$$|\det(\varphi^{-1})'(x)| = [(1-x_1)(1-x_1-x_2)\cdots(1-x_1-\cdots-x_{n-1})]^{-1}.$$

14. Using Problems 11 and 13, prove for the simplex S in Problem 13 that

$$\int_S x_1^{a_1-1}x_2^{a_2-1}\cdots x_n^{a_n-1}\,dx = \frac{\Gamma(a_1)\Gamma(a_2)\cdots\Gamma(a_n)}{\Gamma(a_1+\cdots+a_n+1)}$$

when $a_j > 0$ for all j.

Problems 15–17 concern the Cantor function for the standard Cantor set.

15. Prove that the values of the Cantor function F for the standard Cantor set are independent of the method of defining the approximating functions F_n on the complementary closed intervals as long as F_n is monotone increasing and continuous.
16. Compute $\int_0^1 F(x)\,dx$ if F is the Cantor function for the standard Cantor set.
17. The Stieltjes measure μ corresponding to the Cantor function for the standard Cantor set C is called the **Cantor measure**. The set C consists of the members of $[0, 1]$ that can be expanded in the digits 0, 1, 2 of base 3 without using any 1's. Show, for each n-tuple of 0's and 2's, that μ attaches measure 2^{-n} to the subset of C whose base 3 expansion begins with that n-tuple.

Problems 18–20 introduce the **Poisson integral formula** for the unit disk in $\mathbb{R}^2$. The Poisson kernel was the subject of Problems 27–29 at the end of Chapter I and is given by

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|}e^{in\theta} = \frac{1-r^2}{1-2r\cos\theta+r^2}.$$

Harmonic functions in the unit disk were the subject of Problems 14–15 at the end of Chapter III and also Problems 10–13 at the end of Chapter IV. The present set of problems begins to relate the Poisson kernel to harmonic functions via convolution.

18. If f is in $L^1\big([-\pi,\pi], \frac{1}{2\pi}\,d\theta\big)$, then the **Poisson integral** of f is the function in the unit disk defined in polar coordinates by
$$u(r,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\varphi)P_r(\theta-\varphi)\,d\varphi.$$
If c_n is the n^{th} Fourier coefficient of f, prove that $u(r,\theta) = \sum_{n=-\infty}^{\infty} c_n r^{|n|}e^{in\theta}$, and conclude that u is harmonic in the open unit disk.

19. If p equals 1 or 2 and if f is in $L^p\big([-\pi,\pi], \frac{1}{2\pi}\,d\theta\big)$, prove that the Poisson integral $u(r,\theta)$ of f has the properties that $\|u(r,\,\cdot\,)\|_p \le \|f\|_p$ for $0 \le r < 1$ and that $u(r,\,\cdot\,)$ tends to f in L^p in the sense that $\lim_{r\uparrow 1}\|u(r,\,\cdot\,) - f\|_p = 0$.

20. Suppose that f is in $L^\infty\big([-\pi,\pi], \frac{1}{2\pi}\,d\theta\big)$ and that $u(r,\theta)$ is the Poisson integral of f.
    (a) Prove that $\lim_{r\uparrow 1} u(r,\theta) = f(\theta)$ uniformly on any set of θ's where f is uniformly continuous.
    (b) For f of class C^2, prove that the Poisson integral of f is the only harmonic function $u(r,\theta)$ in the disk such that $\lim_{r\uparrow 1} u(r,\theta) = f(\theta)$ uniformly in θ.
    (c) Prove that $u(r,\,\cdot\,)$ tends to f weak-star in L^∞ relative to L^1 in the sense that $\lim_{r\uparrow 1}\int_{-\pi}^{\pi} u(r,\theta)g(\theta)\,d\theta = \int_{-\pi}^{\pi} f(\theta)g(\theta)\,d\theta$ for all g in L^1. (Weak-star convergence was defined in Section V.9.)

Problems 21–25 concern functions of bounded variation. For such a function f, the positive and negative variations of f were defined in Section 9, and their values at x were denoted by $V^+(f)(x)$ and $V^-(f)(x)$.

21. Prove that the product of two functions of bounded variation on $[a,b]$ is of bounded variation.

22. This problem concerns a certain minimality property of the decomposition $f(x) = f(a) + V^+(f)(x) - V^-(f)(x)$ of a function f of bounded variation on $[a,b]$. Prove that if g_1 and g_2 are any two nonnegative monotone increasing functions such that $f(x) = f(a) + g_1(x) - g_2(x)$ for all x, then $V^+(f)(x) \le g_1(x)$ and $V^-(f)(x) \le g_2(x)$.

23. Prove that if f is of bounded variation on $[a,b]$ and is continuous at a point x in (a,b), then both $V^+(f)$ and $V^-(f)$ are continuous at x.

24. If f is of bounded variation on $[a,b]$, define the **total variation** of f as the function given by
$$V(f)(x) = \sup_{\substack{P \text{ with } x_0=a\\ \text{and } x_n=x}} \sum_{i=1}^{n} \big|f(x_i) - f(x_{i-1})\big|,$$
the supremum being taken over all partitions of $[a,x]$. Prove that $V(f)(x) = V^+(f)(x) + V^-(f)(x)$ for all x.

25. Prove that the function f on $[-1, 1]$ given by

$$f(x) = \begin{cases} x \sin(1/x) & \text{for } x \neq 0, \\ 0 & \text{for } x = 0, \end{cases}$$

is not of bounded variation. Prove or disprove that the function g on $[-1, 1]$ given by

$$g(x) = \begin{cases} x^2 \sin(1/x) & \text{for } x \neq 0, \\ 0 & \text{for } x = 0, \end{cases}$$

is of bounded variation.

CHAPTER VII

Differentiation of Lebesgue Integrals on the Line

Abstract. This chapter concerns the Fundamental Theorem of Calculus for the Lebesgue integral, viewed from Lebesgue's perspective but slightly updated.

Section 1 contains Lebesgue's main tool, a theorem saying that monotone functions on the line are differentiable almost everywhere. A relatively easy consequence is Fubini's theorem that an absolutely convergent series of monotone increasing functions may be differentiated term by term. The result that the indefinite integral $\int_a^x f(t)\,dt$ of a locally integrable function f is differentiable almost everywhere with derivative f follows readily.

Section 2 addresses the converse question of what functions F have the property for a particular f that the integral $\int_a^b f(t)\,dt$ can be evaluated as $F(b)-F(a)$ for all a and b. The development involves a decomposition theorem for monotone increasing functions and a corresponding decomposition theorem for Stieltjes measures. The answer to the converse question when $f \geq 0$ and $F' = f$ almost everywhere is that F is "absolutely continuous" in a sense defined in the section.

1. Differentiation of Monotone Functions

The generalization of the Fundamental Theorem of Calculus to the Lebesgue integral was the crowning achievement of Lebesgue's book. We have already stated and proved a particular result in that direction as Corollary 6.40, using a more recent method that is of continual applicability in analysis. The statement of the part of the Fundamental Theorem in that corollary is that $\int_a^x f(t)\,dt$ is differentiable almost everywhere with derivative $f(x)$ if f is a Borel function on the line that is integrable on every bounded interval.

In this chapter we shall develop that and allied results using something closer to Lebesgue's original method. These allied results are chiefly of historical interest, no longer being of great importance as analytic tools. However, their beauty is undeniable and by itself justifies their inclusion in this book. In addition, these allied results motivate some results in Chapter IX, particularly the Radon–Nikodym Theorem, that might seem strange indeed if the historical background were omitted.

The starting point is the almost-everywhere differentiability of monotone functions on the line, given in Theorem 7.2 below. Since monotone functions include the distribution functions of Stieltjes measures, this differentiability shows at once that functions of the form $\int_a^x f(t)\,dt$ with $f \ge 0$ are differentiable almost everywhere, and then we are well on our way toward a more traditional proof of the Fundamental Theorem for the Lebesgue integral. The advantage of starting with all monotone functions is that one can address at the same time the differentiability of *all* distribution functions of Stieltjes measures, not just those of measures $f(t)\,dt$. From this fact one can attack the question of how close the derivative $f(t)$ is to determining the function of which it is the derivative almost everywhere. This is the second aspect of the traditional Fundamental Theorem of Calculus as in Theorem 1.32: for the case of continuous f, any two functions with derivative f *everywhere* on an interval differ by a constant.

There is a certain formal similarity between the theory of differentiation of monotone functions and the theory of the Hardy–Littlewood Maximal Theorem as in Chapter VI. Wiener's Covering Lemma captured the geometric core of the theorem in Chapter VI, and another covering lemma captures the geometric core here. This is the Rising Sun Lemma, which will be given as Lemma 7.1.

By way of preliminaries, any open subset U of $\mathbb{R}^1$ is uniquely the union of countably many disjoint open intervals, the open interval containing a point x in U being the union of all connected subsets of U containing x. These sets give the required decomposition of U by Propositions 2.48 and 2.51. An open subset of an interval (a, b) is necessarily open in $\mathbb{R}^1$, and hence it too is uniquely the countable union of disjoint open intervals.

Lemma 7.1 (Rising Sun Lemma).[1] Let $g : [a, b] \to \mathbb{R}$ be continuous, and define

$$E = \big\{x \in (a, b) \bigm| \text{there exists } \xi \in (a, b) \text{ with } \xi > x \text{ and } g(\xi) > g(x)\big\}.$$

The set E is open in (a, b). If E is written as the disjoint union of open intervals with endpoints a_k and b_k, then $g(a_k) \le g(b_k)$ for each k.

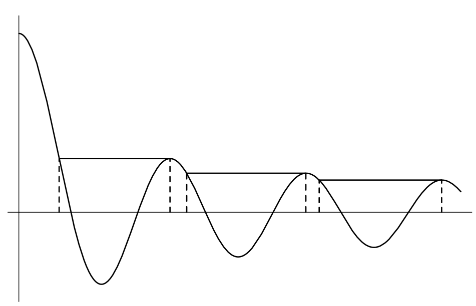


FIGURE 7.1. Rising Sun Lemma. Graph showing three open intervals produced by the lemma.

[1] Some authors call this result Riesz's Lemma.

REMARK. The Rising Sun Lemma is so named because of the situation in Figure 7.1. The sun rises in the east, viewed as the direction of the positive x axis. It casts shadows within the graph of g, and the content of the lemma is the nature of those shadows. Although the conclusion of the lemma is that $g(a_k) \le g(b_k)$ for all k, the reader can observe in the figure that $g(a_k) = g(b_k)$ for the open intervals that are shown. This observation is valid in general except possibly when $a_k = a$, but the observation is not needed in the proof of Theorem 7.2 below.

PROOF. If $x_0 \in E$ and $\xi \in (a, b)$ have $\xi > x_0$ and $g(\xi) > g(x_0)$, then every x in (a, ξ) with $|g(x) - g(x_0)| < \frac{1}{2}(g(\xi) - g(x_0))$ lies in E. Hence E is open.

Let E be the disjoint union of intervals (a_k, b_k). Fix attention on one such interval (a_k, b_k). We make critical use of the fact that the point b_k is not in E. If x_0 satisfies $a_k < x_0 < b_k$, we prove that $g(x_0) \le g(b_k)$. Once we do so, we can let x_0 decrease to a_k and use continuity to obtain the assertion $g(a_k) \le g(b_k)$ of the lemma.

Arguing by contradiction, suppose that $g(x_0) > g(b_k)$. Since x_0 is in E, there exists $x_1 > x_0$ with $g(x_1) > g(x_0)$. If $x_1 > b_k$, then the inequality $g(x_1) > g(x_0) > g(b_k)$ forces b_k to be in E. Since b_k is not in E, we conclude that $x_1 \le b_k$. The set of all x with $x_1 \le x \le b_k$ and $g(x) \ge g(x_1)$ is closed, bounded, and nonempty, and we let x_2 be its largest element, so that $x_2 \le b_k$.

Since $g(x_2) \ge g(x_1) > g(x_0) > g(b_k)$, we must have $x_2 < b_k$; in fact, $x_2 = b_k$ would yield the contradiction $g(b_k) > g(b_k)$. From $a_k < x_0 < x_2 < b_k$ and $(a_k, b_k) \subseteq E$, we see that x_2 is in E. Hence there is some $\xi > x_2$ with $g(\xi) > g(x_2)$. Then the conditions $g(\xi) > g(b_k)$ and $b_k \notin E$ together force ξ to be $\le b_k$. So $x_2 < \xi \le b_k$ with $g(\xi) \ge g(x_1)$, in contradiction to the maximality of x_2. This contradiction allows us to conclude that $g(x_0) \le g(b_k)$, and the proof is complete.

Theorem 7.2 (Lebesgue). If F is a monotone increasing function on an interval, then F is differentiable almost everywhere in this sense: the set where F fails to be differentiable is a Lebesgue measurable set of Lebesgue measure 0. In addition, if the definition of F' is extended so that $F'(x) = 0$ at every point where F is not differentiable, then F' is Lebesgue measurable.

REMARK. Recall that any monotone increasing function F can have only countably many discontinuities, and these are all given by jumps. In other words, F has, at each point x, left and right limits $F(x-)$ and $F(x+)$, and the only possible discontinuities occur when one or both of the equalities $F(x-) = F(x)$ and $F(x) = F(x+)$ fail.

PROOF. The second statement is a consequence of the first. In fact, if E is the Lebesgue measurable set of measure 0 where F is nondifferentiable and if B is a Borel set of measure 0 containing E, then the sequence of Borel functions

$G_n(x) = \frac{1}{1/n}(F(x+1/n) - F(x))$ converges everywhere on B^c to a function G. If G is extended to the domain of F by defining it to be 0 on B, then G is a Borel function that equals F' except on a subset of B, and hence F' is Lebesgue measurable.

Let us come to the conclusion about differentiability. Possibly by taking the union of countably many sets, we may assume that the domain of F is a bounded interval $[a, b]$. For $a < x < b$, define

$$U_r(x) = \limsup_{h \downarrow 0} \tfrac{1}{h}(F(x+h) - F(x))$$

and

$$L_r(x) = \liminf_{h \downarrow 0} \tfrac{1}{h}(F(x+h) - F(x)),$$

$$U_l(x) = \limsup_{h \uparrow 0} \tfrac{1}{h}(F(x+h) - F(x))$$

and

$$L_l(x) = \liminf_{h \uparrow 0} \tfrac{1}{h}(F(x+h) - F(x)).$$

We shall prove that

$$U_r(x) < +\infty$$

and

$$U_r(x) \leq L_l(x)$$

almost everywhere. If the latter inequality is applied to $-F(-x)$, we obtain also

$$U_l(x) \leq L_r(x)$$

almost everywhere. Putting these inequalities together, we have $U_l(x) \leq L_r(x) \leq U_r(x) \leq L_l(x) \leq U_l(x)$ almost everywhere, and equality must hold throughout, almost everywhere. The points where equality holds throughout and also $U_r(x) < +\infty$ are the points where F is differentiable, and hence the two inequalities $U_r(x) < +\infty$ and $U_r(x) \leq L_l(x)$ prove the theorem.

For most of the proof, we shall assume that F is continuous. At the end we return and show how to modify the proof to handle discontinuous F. First we consider the inequality $U_r(x) < +\infty$. The subset E of (a, b) where this inequality fails is, for each positive integer n, contained in the set where $U_r(x) > n$. If $U_r(x) > n$, then $\dfrac{F(\xi) - F(x)}{\xi - x} > n$ for some $\xi > x$. That is, $g(\xi) > g(x)$ for the continuous function $g(x) = F(x) - nx$. In the notation of Lemma 7.1, E is covered by a system of disjoint open intervals (a_k, b_k) such that $g(a_k) \leq g(b_k)$ for each such interval. Thus $n(b_k - a_k) \leq F(b_k) - F(a_k)$ for each. Summing on k gives $n \sum_k (b_k - a_k) \leq \sum_k (F(b_k) - F(a_k)) \leq F(b) - F(a)$. Thus the exceptional set E can be covered by a system of open intervals of total measure $\leq \frac{1}{n}(F(b) - F(a))$. Since n is arbitrary, Proposition 5.39 shows that E is Lebesgue measurable of Lebesgue measure 0.

Next we prove that $U_r(x) \le L_l(x)$ almost everywhere on (a, b). If $0 \le p < q$ are rational numbers, we prove that the set E_{pq} where

$$L_l(x) < p < q < U_r(x)$$

has Lebesgue measure 0. The countable union of such sets is the exceptional set in question, and thus we will have proved that the exceptional set has measure 0. If $L_l(x) < p$, then there exists $\xi \in (a, b)$ with $\xi < x$ and $\dfrac{F(\xi) - F(x)}{\xi - x} < p$, hence with $p\xi - F(\xi) < px - F(x)$. Define $g(z) = pz + F(-z)$ for z in $[-b, -a]$. If $y = -x$ and $\eta = -\xi$, then $p\eta + F(-\eta) > py + F(-y)$ and hence $g(\eta) > g(y)$ with $\eta > y$. Applying Lemma 7.1 to g on the interval $[-b, -a]$, we obtain a disjoint system of open intervals $(-b_i, -a_i)$ covering the set of y's where $L_l(-y) < p$ and having $g(-b_i) \le g(-a_i)$ in each case. Thus $-pb_i + F(b_i) \le -pa_i + F(a_i)$. In other words, the set of x's where $L_l(x) < p$ is covered by a disjoint system of open intervals (a_i, b_i) such that

$$F(b_i) - F(a_i) \le p(b_i - a_i) \qquad (*)$$

for each such interval. Applying the lemma to $g_p(x) = F(x) - qx$ on the interval $[a_i, b_i]$, we obtain a disjoint system of open intervals (a_{ij}, b_{ij}) indexed by j and having $g_p(a_{ij}) \le g_p(b_{ij})$. Thus $(*)$ and

$$q(b_{ij} - a_{ij}) \le F(b_{ij}) - F(a_{ij}) \qquad (**)$$

hold in each case. Summing $(**)$ over j, we obtain

$$q \sum_j (b_{ij} - a_{ij}) \le \sum_j (F(b_{ij}) - F(a_{ij})) \le F(b_i) - F(a_i) \le p(b_i - a_i). \quad (\dagger)$$

Summing this inequality over i and dividing by q gives

$$m(E_{pq}) \le \sum_{i,j} (b_{ij} - a_{ij}) \le (p/q)(b - a).$$

If we repeat this argument with $[a_{ij}, b_{ij}]$ in place of $[a, b]$, we obtain intervals (a_{ijuv}, b_{ijuv}) and an inequality

$$m(E_{pq}) \le \sum_{i,j,u,v} (b_{ijuv} - a_{ijuv}) \le (p/q) \sum_{i,j} (b_{ij} - a_{ij}) \le (p/q)^2(b - a).$$

Iteration gives $m(E_{pq}) \le (p/q)^n(b - a)$ for every n, and therefore $m(E_{pq}) = 0$. This completes the proof in the case that F is continuous.

If F is possibly discontinuous, we modify Lemma 7.1 and the present proof as follows. Each function g that arises has right and left limits $g(x+)$ and $g(x-)$ at each point x, and we let $G(x)$ be the largest of $g(x-)$, $g(x)$, and $g(x+)$. A modified Lemma 7.1 says that the set of x in (a, b) for which there is some $\xi \in (a, b)$ with $\xi > x$ and $g(\xi) > G(x)$ is an open set whose component intervals (a_k, b_k) have $g(a_k+) \leq G(b_k)$ for each k. Going over the proof of Lemma 7.1 carefully and changing g to G as necessary, we obtain a proof of the modified Lemma 7.1.

The modifications necessary to the present proof are as follows. In the proof that $U_r(x) < +\infty$ almost everywhere, the set E is to be taken to be the set where F is continuous and this inequality fails. The inequality that results from applying the modified Lemma 7.1 is $n(b_k - a_k) \leq F(b_k+) - F(a_k+)$, and this inequality can be summed on k without any further change. Similarly in the proof that $U_r(x) \leq L_l(x)$ almost everywhere, the set E_{pq} is to be taken to be the set where F is continuous and $L_l(x) < p < q < U_r(x)$. Inequality $(*)$ becomes $F(b_i-) - F(a_i+) \leq p(b_i - a_i)$. When we consider the interval $[a_i, b_i]$, the value of $F(b_i+)$ is not relevant, and the value of $F(b_i)$ can be adjusted to equal $F(b_i-)$ for purposes of understanding F between a_i and b_i. With that understanding, inequality $(**)$ becomes $q(b_{ij} - a_{ij}) \leq F(b_{ij}+) - F(a_{ij}+)$, and step $(\dagger)$ is replaced by

$$q \sum_j (b_{ij} - a_{ij}) \leq \sum_j (F(b_{ij}+) - F(a_{ij}+)) \leq F(b_i-) - F(a_i+) \leq p(b_i - a_i).$$

The two inequalities at the ends have come about from $(*)$ and $(**)$, and the critical observation is that the convention $F(b_i) = F(b_i-)$ makes the middle inequality hold. The rest of the argument proceeds as in the case that F is continuous, and then the theorem is completely proved.

Theorem 7.3 (Fubini's theorem on the differentiation of series of monotone functions). If $F = \sum F_n$ is an absolutely convergent sequence of monotone increasing functions on $[a, b]$, then $F'(x) = \sum_{n=1}^\infty F_n'(x)$ almost everywhere.

PROOF. Without loss of generality, we may assume that $F_n(a) = 0$ for all n. Then $F_n(x) \geq 0$ for all n and x. Possibly by lumping terms, we may assume also that $F(b) - \sum_{k=1}^n F_k(b) \leq 2^{-n}$. Since $F(x) - \sum_{k=1}^n F_k(x)$ is a monotone increasing function that is 0 for $x = a$, we have

$$0 \leq F(x) - \sum_{k=1}^n F_k(x) \leq 2^{-n} \qquad (*)$$

for $a \leq x \leq b$. The decomposition $F(x) = \sum_{k=1}^n F_k(x) + \big(\sum_{k=n+1}^\infty F_k(x)\big)$ exhibits F as the sum of $n+1$ monotone increasing functions, and thus we have

$\sum_{k=1}^n F_k'(x) \le F'(x)$ at all points where all the derivatives exist. In view of Theorem 7.2, this inequality holds almost everywhere. Consequently

$$0 \le \sum_{k=1}^{\infty} F'(k) \le F'(x) \tag{$**$}$$

almost everywhere. Now consider the series

$$G(x) = \sum_{n=1}^{\infty} \Big(F(x) - \sum_{k=1}^{n} F_k(x) \Big).$$

Then

$$0 \le G(x) - \sum_{n=1}^{N} \Big(F(x) - \sum_{k=1}^{n} F_k(x) \Big) \le \sum_{n=N+1}^{\infty} 2^{-n} = 2^{-N}.$$

Thus G satisfies the same kind of inequality that F did in $(*)$, and we can conclude that G satisfies the analog of $(**)$, namely

$$0 \le \sum_{n=1}^{\infty} \Big(F'(x) - \sum_{k=1}^{n} F_k'(x) \Big) \le G'(x).$$

The right side is finite almost everywhere by Theorem 7.2, and thus the individual terms $F'(x) - \sum_{k=1}^n F_k'(x)$ of the series tend to 0 almost everywhere. This completes the proof.

From Theorems 7.2 and 7.3, we can derive the first part of Lebesgue's form of the Fundamental Theorem of Calculus. This same result was stated as Corollary 6.40 and was proved in Chapter VI by using the Hardy–Littlewood Maximal Theorem.

Corollary 7.4 (first part of Lebesgue's form of the Fundamental Theorem of Calculus). If f is integrable on every bounded subset of $\mathbb{R}^1$, then $\int_a^x f(y)\,dy$ is differentiable almost everywhere and

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x) \qquad \text{almost everywhere.}$$

PROOF. It is enough to prove the theorem for functions vanishing off an interval $[a, b]$. Let $\mathcal{A}$ be the set of all Borel sets $E \subseteq [a, b]$ such that $\frac{d}{dx}\int_a^x I_E(t)\,dt = I_E(x)$ almost everywhere. Then $\mathcal{A}$ contains the elementary sets within $[a, b]$, and $\mathcal{A}$ is closed under complements within $[a, b]$. If $\{E_n\}$ is an increasing sequence of sets in $\mathcal{A}$ with $E_0 = \varnothing$ and with union E, let us write $I_E = \sum_{n=1}^{\infty} (I_{E_n} - I_{E_{n-1}})$. This is a series of nonnegative functions. Putting $F(x) = \int_a^x I_E(t)\,dt$ and $F_n(x) = \int_a^x (I_{E_n}(t) - I_{E_{n-1}}(t))\,dt$ and applying Corollary 5.27, we obtain $F(x) = \sum_{n=1}^{\infty} F_n(x)$. Then Theorem 7.3 gives $F'(x) = \sum_{n=1}^{\infty} F_n'(x) = \lim_N \sum_{n=1}^{N} F_n'(x) = \lim_N \sum_{n=1}^{N} (I_{E_n}(x) - I_{E_{n-1}}(x)) = \lim_N I_{E_N}(x) = I_E(x)$ almost everywhere. Thus E is in $\mathcal{A}$, and $\mathcal{A}$ is closed under increasing countable unions. Since $\mathcal{A}$ is closed under complements as well, $\mathcal{A}$ is closed under decreasing countable intersections. Then the Monotone Class Lemma (Lemma 5.43) shows that $\mathcal{A}$ contains all Borel sets.

Now consider the set $\mathcal{F}$ of all integrable Borel functions f for which the almost-everywhere equality $\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$ holds. We have just seen that $\mathcal{F}$ contains all indicator functions of Borel subsets of $[a, b]$. By linearity, $\mathcal{F}$ contains all nonnegative simple functions vanishing off $[a, b]$. Let $f \geq 0$ be an integrable function on $[a, b]$, and let $\{s_n\}$ be an increasing sequence of nonnegative simple functions with pointwise limit f. The functions s_n are in $\mathcal{F}$. Put $s_0 = 0$, and let $F(x) = \int_a^x f(t)\,dt$ and $F_n(x) = \int_a^x (s_n(t) - s_{n-1}(t))\,dt$. Since $s_n \geq s_{n-1}$, each F_n is monotone increasing. Corollary 5.27 shows that $F(x) = \sum_{n=1}^{\infty} F_n(x)$, and Theorem 7.3 then shows that $F'(x) = \sum_{n=1}^{\infty} F_n'(x) = \lim_N \sum_{n=1}^{N} F_n'(x) = \lim_N \sum_{n=1}^{N} (s_n(x) - s_{n-1}(x)) = \lim_N s_n(x) = f(x)$ almost everywhere. Thus $\mathcal{F}$ contains all nonnegative integrable Borel functions, and by linearity it contains all integrable Borel functions.

2. Absolute Continuity, Singular Measures, and Lebesgue Decomposition

In this section we address questions about the Lebesgue integral raised by the second part of the Fundamental Theorem of Calculus in Theorem 1.32. For continuous integrands f, the result is a kind of uniqueness statement, asserting that any function with derivative f differs from $\int_a^x f(t)\,dt$ by a constant function. From a practical point of view, this is the really important part of the theorem for calculus, since it provides a technique for evaluating definite integrals: find any function whose derivative is the given function, evaluate it at the endpoints, and subtract the results. With the Lebesgue integral and equality of derivatives only almost everywhere, the uniqueness result is not as sharp. The practical aspect of a uniqueness theorem is largely lost, and the resulting theory ends up having to be appreciated only as an end in itself. We begin at the following point.

Proposition 7.5. Every monotone increasing function on $\mathbb{R}^1$ is uniquely the sum of an indefinite integral $G(x) = \int_0^x f(t)\,dt$, where $f \geq 0$ is integrable on every bounded interval, and a monotone increasing function H such that $H'(x) = 0$ almost everywhere.

PROOF. Let F be a given monotone increasing function on $\mathbb{R}^1$. If $F = G + H$ with G as in the statement of the proposition and with $H'(x) = 0$ almost everywhere, Corollary 7.4 shows that we must have $f = F'$. This proves uniqueness.

For existence we take $f = F'$. Regard h as a positive number tending to 0 through some sequence, so that $h^{-1}(F(t+h) - F(t))$ tends to $f'(t)$ for almost every t. If $a < b$, then

$$\begin{aligned}\int_a^b \frac{F(t+h) - F(t)}{h}\,dt &= \frac{1}{h}\int_{a+h}^{b+h} F(t)\,dt - \frac{1}{h}\int_a^b F(t)\,dt \\ &= \frac{1}{h}\int_b^{b+h} F(t)\,dt - \frac{1}{h}\int_a^{a+h} F(t)\,dt.\end{aligned}$$

The right side tends to $F(b) - F(a)$ if a and b are points of continuity of F. By Fatou's Lemma (Theorem 5.29), $\int_a^b f(t)\,dt \le F(b) - F(a)$ if a and b are points of continuity of F. The points of continuity of F are dense, and thus for general a and b, we can find sequences of points of continuity decreasing to a and increasing to b. Passing to the limit, we obtain

$$\int_a^b f(t)\,dt \le F(b-) - F(a+) \le F(b) - F(a) \qquad (*)$$

for all a and b. Hence f is integrable. With $G(x)$ as in the statement of the proposition, $(*)$ gives $G(b) - G(a) \le F(b) - F(a)$. Equivalently, $F(a) - G(a) \le F(b) - G(b)$. Thus the function $H(x) = F(x) - G(x)$ is monotone increasing with $F = G + H$. Since F and G have derivative f almost everywhere, H has derivative 0 almost everywhere.

Thus we want to identify all monotone increasing functions with derivative zero almost everywhere. The first step is to see that the question of discontinuities of a monotone function can be completely eliminated from the problem.

Proposition 7.6. Let c be a real number. If $\{x_n\}$ is a sequence in $[a, b]$ and if $\{c_n\}$ and $\{d_n\}$ are sequences of positive real numbers with $\sum c_n$ finite and $\sum d_n$ finite, then the function

$$F(x) = c + \sum_{\substack{n \text{ with} \\ x_n \le x}} c_n + \sum_{\substack{n \text{ with} \\ x_n < x}} d_n$$

is a monotone increasing function on $[a, b]$ with $F'(x) = 0$ almost everywhere.

PROOF. The function F is certainly monotone increasing. It is the convergent sum of the constant function c and monotone increasing functions of the form

$$F_n(x) = \begin{cases} 0 & \text{for } x < x_n, \\ c_n & \text{for } x = x_n, \\ c_n + d_n & \text{for } x > x_n, \end{cases}$$

and the function F_n has derivative 0 except at the point x_n. Thus the proposition follows immediately from Theorem 7.3.

A monotone increasing function on the line whose restriction to every closed bounded interval is of the form in Proposition 7.6 is called a **saltus function**; the name comes from the Latin word for "jump." Since $\mathbb{R}^1$ is the countable union of closed bounded intervals, it follows from Proposition 7.6 that every saltus function has derivative 0 almost everywhere.

Proposition 7.7. Any monotone increasing function F on $\mathbb{R}^1$ is uniquely the sum $F = G + S$ of a continuous monotone increasing function G with $G(0) = 0$ and a saltus function S.

PROOF. For existence, it is enough to obtain the decomposition without insisting on the normalization $G(0) = 0$, since the sum of a saltus function and a constant is a saltus function. Let x_0 be a point of continuity of F, and enumerate the points of discontinuity of F as x_n, $n \geq 1$. For each $n \geq 1$, define $c_n = F(x_n) - F(x_n-)$ and $d_n = F(x_n+) - F(x_n)$. Let S be the saltus function

$$S(x) = \begin{cases} \sum_{x_0 \leq x_n \leq x} c_n + \sum_{x_0 \leq x_n < x} d_n & \text{if } x \geq 0, \\ -\sum_{x < x_n \leq x_0} c_n - \sum_{x \leq x_n < x \leq x_0} d_n & \text{if } x \leq 0, \end{cases}$$

and put $G = F - S$. Then G is continuous everywhere. To see that G is monotone increasing, let $a < b$ be points of continuity of F and S. We start from the equality $S(x_n+) - S(x_n-) = F(x_n+) - F(x_n-)$ and sum for x_n with $a < x_n < b$ to obtain

$$\begin{aligned} S(b) - S(a) &= \sum_{a < x_n < b} (S(x_n+) - S(x_n-)) \\ &= \sum_{a < x_n < b} (F(x_n+) - F(x_n-)) \\ &\leq F(b) - F(a). \end{aligned}$$

Hence $F(a) - S(a) \leq F(b) - S(b)$, and we conclude that $G(a) \leq G(b)$ at all points of continuity $a < b$ of F and S. These points are dense, and G is continuous everywhere. Hence $G(a) \leq G(b)$ whenever $a < b$, and G is monotone increasing. This proves existence. Uniqueness follows from the fact that $S(b-) - S(a+) = \sum_{a < x_n < b} (F(x_n+) - F(x_n-))$ whenever $a < b$, and the proof is complete.

Consequently we need to understand the continuous monotone increasing functions F on the line with $F'(x) = 0$ almost everywhere. The Cantor function for the standard Cantor set, constructed as in Section VI.8, is an example. For such a function, $F - F(0)$ satisfies the defining properties of the distribution function of a Stieltjes measure μ on $\mathbb{R}^1$. The continuity of F is equivalent to the fact that μ contains no point masses. The following property isolates the meaning of having derivative zero almost everywhere.

Proposition 7.8. Suppose that μ is a Stieltjes measure with no point masses. If the distribution function F of μ has $F'(x) = 0$ at every point of a Borel set E, then $\mu(E) = 0$.

REMARK. The proof will use the Rising Sun Lemma (Lemma 7.1). Problem 3 at the end of the chapter asks for an alternative proof by means of Wiener's Covering Lemma (Lemma 6.41). A proof using Wiener's Covering Lemma does not make use of the continuity of F, and therefore it is not necessary to assume in the proposition that μ has no point masses.

PROOF. We may confine our attention to an interval $[a, b]$, taking E to be a subset of $[a, b]$. Since μ has no point masses, we may discard a and b from E. Fix a positive integer n. For every point x in E, we have $F'(x) < \frac{1}{n}$. Therefore to each such x, we can associate some $\xi > x$ with ξ in (a, b) such that

$$\frac{F(\xi) - F(x)}{\xi - x} < \frac{1}{n}.$$

This inequality says that $\frac{1}{n}\xi - F(\xi) > \frac{1}{n}x - F(x)$, hence that the continuous function g with $g(x) = \frac{1}{n}x - F(x)$ has $g(\xi) > g(x)$. The Rising Sun Lemma (Lemma 7.1) applies and shows that E is covered by countably many disjoint open intervals (a_k, b_k) with $g(a_k) \le g(b_k)$. Thus $\frac{1}{n}a_k - F(a_k) \le \frac{1}{n}b_k - F(b_k)$ for each k. Adding, we obtain

$$\mu(E) \le \sum_k \mu((a_k, b_k)) = \sum_k F(b_k) - F(a_k)) \le \tfrac{1}{n} \sum_k (b_k - a_k) \le \tfrac{1}{n}(b - a).$$

Since n is arbitrary, $\mu(E) = 0$.

Again consider a continuous monotone function F with derivative zero almost everywhere. The function $F - F(0)$ is the distribution function of some Stieltjes measure μ with no point masses, and Proposition 7.8 shows that there is a Borel set E such that $\mu(E) = 0$ and $m(E^c) = 0$, where m is Lebesgue measure. In other words, μ is concentrated completely on the set E^c of Lebesgue measure 0. A Stieltjes measure μ for which there is a Borel set F with $\mu(F^c) = 0$ and

$m(F) = 0$ is called a **singular** Stieltjes measure or a "Stieltjes measure singular with respect to Lebesgue measure." If also it contains no point masses, it is said to be **continuous singular**. The Stieltjes measure associated to the Cantor function for the standard Cantor set is an example. We can summarize matters either in terms of decompositions of monotone functions or in terms of decompositions of Stieltjes measures. The result in the case of monotone functions is a first answer to the question of uniqueness in the Fundamental Theorem of Calculus for the Lebesgue integral; the result in the case of Stieltjes measures gives the **Lebesgue decomposition** of Stieltjes measures.

Theorem 7.9. Every monotone increasing function F on $\mathbb{R}^1$ decomposes uniquely as the sum $F = G + H + S$, where G is the indefinite integral $G(x) = \int_0^x f(t)\,dt$ of a function $f \geq 0$ integrable on every bounded interval, H is the distribution function of a continuous singular measure, and S is a saltus function. The function f is the derivative of F.

PROOF. Proposition 7.7 allows us to write $F = P + S$ uniquely, where S is a saltus function and P is continuous and monotone increasing with $P(0) = 0$. Proposition 7.5 says that $P = G + H$ uniquely, where G is an indefinite integral $G(x) = \int_0^x f(t)\,dt$ and H is monotone increasing with $H'(x) = 0$ almost everywhere. The function f is the derivative of F. The function H has $H(0) = 0$ and is continuous because P and G have these properties, and therefore H is the distribution function of a Stieltjes measure μ containing no point masses. Since $H'(x) = 0$ almost everywhere, Proposition 7.8 shows that μ is singular.

Corollary 7.10 (Lebesgue decomposition). Every Stieltjes measure μ decomposes uniquely as the sum $\mu = f\,dx + \mu_s + \mu_d$, where $f \geq 0$ is a function integrable on every bounded interval, μ_s is a continuous singular measure, and μ_d is a countable sum of point masses such that the sum of the weights on any bounded interval is finite.

PROOF. This follows by applying Theorem 7.9 to the distribution function of μ.

The final question that we address in this section is how to recognize the particular monotone function $G(x) = \int_0^x f(t)\,dt$ from among all the monotone functions $F = G + H + S$ described in Theorem 7.9.

Proposition 7.11. With m denoting Lebesgue measure, the following conditions on a Stieltjes measure μ_a are equivalent:

(a) μ_a is of the form $\mu_a = f\,dx$ for some function $f \geq 0$ that is integrable on every bounded interval,

(b) for each bounded interval $[a, b]$ and number $\epsilon > 0$, there exists a number $\delta > 0$ such that $\mu_a(E) < \epsilon$ whenever E is a Borel subset of $[a, b]$ with $m(E) < \delta$,

(c) $\mu_a(E) = 0$ whenever E is a Borel subset of $\mathbb{R}^1$ with $m(E) = 0$.

REMARK. A Stieltjes measure μ_a satisfying the equivalent conditions in this proposition is said to be **absolutely continuous** or "absolutely continuous with respect to Lebesgue measure." From any of these defining conditions, we see right away that an absolutely continuous measure contains no point masses.

PROOF. Corollary 5.24 shows immediately that (a) implies (b). To see that (b) implies (c), let E be a Borel set E in $\mathbb{R}^1$ with $m(E) = 0$. Applying (b) to $E \cap [a, b]$ gives $\mu_a(E \cap [a, b]) < \epsilon$ for every positive ϵ, and hence $\mu_a(E \cap [a, b]) = 0$. Since $[a, b]$ is arbitrary and μ_a is completely additive, $\mu_a(E) = 0$.

To see that (c) implies (a), we appeal to Corollary 7.10 to decompose μ_a according to the Lebesgue decomposition as

$$\mu_a = f\,dx + \mu_s + \mu_d, \tag{$*$}$$

where μ_s is continuous singular and μ_d is discrete. The measures μ_s and μ_d have the property that there is a Borel set E with $m(E) = 0$ such that $\mu_s(E^c) = \mu_d(E^c) = 0$. Condition (c) shows that $\mu_a(E) = 0$. Evaluating $(*)$ at E, we obtain $0 = \mu_a(E) = 0 + \mu_s(E) + \mu_d(E)$. Therefore $\mu_s(E) = \mu_d(E) = 0$. Since $\mu_s(E^c) = \mu_d(E^c) = 0$ also, we must have $\mu_s = \mu_d = 0$, and then $(*)$ shows that $\mu_a = f\,dx$.

In Chapter IX the implication (c) implies (a) will be generalized to a result in abstract measure theory known as the Radon–Nikodym Theorem. Meanwhile, it is conditions (b) and (c) that we can translate into a condition on the corresponding distribution function, and then we shall have our second and final answer to the question of uniqueness in the Fundamental Theorem of Calculus for the Lebesgue integral. A monotone increasing function F on the line is said to be **absolutely continuous** if for each bounded interval $[a, b]$ and number $\epsilon > 0$, there exists a $\delta > 0$ such that on any countable disjoint union $\bigcup_k (a_k, b_k)$ of intervals within $[a, b]$ having total length $< \delta$, the variation $\sum_k (F(b_k) - F(a_k))$ of F on that union of intervals is $< \epsilon$.

Proposition 7.12. A Stieltjes measure is absolutely continuous if and only if its distribution function is absolutely continuous.

PROOF. Let μ be a Stieltjes measure with distribution function F. Suppose that μ is absolutely continuous. Fix an interval $[a, b]$, let $\epsilon > 0$ be given, and choose $\delta > 0$ by (b) in Proposition 7.11 such that $m(E) < \delta$ implies $\mu(E) < \epsilon$.

If the set $A = \bigcup_k (a_k, b_k)$ is a countable disjoint union of intervals within $[a, b]$ having total length $< \delta$, then $m(A) < \delta$, and hence $\mu(A) < \epsilon$. Therefore $\sum_k (F(b_k) - F(a_k)) = \sum_k \mu((b_k - a_k)) = \mu(A) < \epsilon$, and we conclude that F is absolutely continuous.

Conversely suppose that F is absolutely continuous, and suppose that E is a Borel set with $m(E) = 0$. Fix an interval $[a, b]$, and let $\epsilon > 0$ be given. By absolute continuity of F, there exists a $\delta > 0$ such that on any countable disjoint union $\bigcup_k (a_k, b_k)$ of intervals within $[a, b]$ having total length $< \delta$, the variation $\sum_k (F(b_k) - F(a_k))$ of F on that union of intervals is $< \epsilon$. With δ defined in this way, we can find a countable disjoint union of intervals $\bigcup_k (a_k, b_k)$ covering $E \cap [a, b]$ and having total length $< \delta$. Then $\mu(E \cap [a, b]) \le \mu\big(\bigcup_k (a_k, b_k)\big) = \sum_k \mu((a_k, b_k)) = \sum_k ((F(b_k) - F(a_k)) < \epsilon$. Since ϵ is arbitrary, $\mu(E \cap [a, b]) = 0$. Since $[a, b]$ is arbitrary, $\mu(E) = 0$. Therefore μ satisfies (c) in Proposition 7.11 and is absolutely continuous.

Corollary 7.13 (second part of Lebesgue's form of the Fundamental Theorem of Calculus). Let F be a monotone increasing function on $\mathbb{R}^1$, and let f be its almost-everywhere derivative. Then $\int_a^b f(t)\,dt = F(b) - F(a)$ whenever $a < b$ if and only if F is absolutely continuous.

PROOF. By Theorem 7.9 we can write $F(x) = \int_0^x f(t)\,dt + H(x) + S(x)$, where H is the distribution function of a continuous singular measure and S is a saltus function. For $a < b$, we then have

$$F(b) - F(a) = \int_a^b f(t)\,dt + (H(b) - H(a)) + (S(b) - S(a)).$$

If $F(b) - F(a) = \int_a^b f(t)\,dt$ whenever $a < b$, then the monotonicity of H and S forces H and S to be constant functions, say with $H(0) + S(0) = c$. Substituting, we see that $F(x) = \int_0^x f(t)\,dt + c$ for all x. The function $\int_0^x f(t)\,dt$ is absolutely continuous by Proposition 7.12, and the additive constant c does not hurt matters. Thus F is absolutely continuous.

Conversely if F is absolutely continuous, then it is continuous, and its monotonicity forces $F - F(0)$ to be a distribution function of some Stieltjes measure μ. Proposition 7.12 shows that the measure μ is absolutely continuous, and Proposition 7.11 shows that μ is of the form $\mu = g\,dx$. Therefore $F(x) - F(0) = \int_0^x g(t)\,dt$. By Corollary 7.4, $g = F' = f$. Hence $F(b) - F(a) = \int_a^b f(t)\,dt$ whenever $a < b$.

3. Problems

1. In the Rising Sun Lemma (Lemma 7.1), show that $g(a_k) = g(b_k)$ if $a_k \ne a$. Give an example of a continuous g for which one of the intervals (a_k, b_k) has $a_k = a$ and $g(a_k) < g(b_k)$.

2. Let m be Lebesgue measure. Does there exist a Lebesgue measurable set E such that $m(E \cap I) = \frac{1}{2}m(I)$ for every bounded interval I? Why or why not?
3. Prove Proposition 7.8 using Wiener's Covering Lemma (Lemma 6.41) instead of the Rising Sun Lemma (Lemma 7.1).
4. Find all continuous monotone increasing functions on $\mathbb{R}^1$ with derivative 0 at all but countably many points.
5. Cantor sets within [0, 1] were introduced in Section II.9. Each is associated to a sequence $\{r_n\}_{n\geq 1}$ of numbers with $0 < r_n < 1$, the standard Cantor set being obtained when $r_n = 1/3$ for every n. Section VI.8 showed how to associate a distribution function to the standard Cantor set, and in similar fashion one can associate a distribution function to any Cantor set. Let C be a Cantor set, let F be the associated distribution function, and let μ be the associated Stieltjes measure. The Lebesgue measure of C is the number $P = \prod_{n=1}^{\infty}(1 - r_n)$. Prove that
   (a) μ is singular if $P = 0$,
   (b) μ is absolutely continuous if $P > 0$, being of the form $P^{-1}I_C(x)\,dx$.

Problems 6–7 concern the **Lebesgue set** of an integrable function f on an interval $[a, b]$. This is the set where $\frac{d}{dx}\int_a^x |f(t) - f(x)|\,dt$ exists and equals 0. Many almost-everywhere convergence results involving f are valid at *every* point of the Lebesgue set. Such results may be regarded as relatively straightforward consequences of Corollary 7.4. Conversely an almost-everywhere convergence theorem that fails to hold at some point of the Lebesgue set might well be expected to involve some new idea.

6. For f integrable on $[a, b]$, prove that almost every point of (a, b) is in the Lebesgue set of f by showing that the Lebesgue set of f is the same as the set where $\frac{d}{dx}\int_a^x |f(t) - r|\,dt \neq |f(x) - r|$ for some rational r.
7. The Fejér kernel, which was defined in Section I.10 and studied further in Section VI.7, is the periodic function defined for t in $[-\pi, \pi]$ by $K_N(t) = \frac{1}{N+1}\frac{1-\cos(N+1)t}{1-\cos t}$. Let f be integrable on $[-\pi, \pi]$, regard f as periodic, and let x be in the Lebesgue set of f. Prove that $\lim_N (K_N * f)(x) = f(x)$ by following these steps:
   (a) Check that the estimates $K_N(t) \leq N + 1$ and $K_N(t) \leq c/(Nt^2)$ are valid for all N and for $|t| \leq \pi$.
   (b) Check that the problem is to show that $\int_{|t|\leq\pi} K_N(t)|f(x - t) - f(x)|\,dt$ tends to 0 as N tends to infinity.
   (c) Break the integral in (b) into pieces where $|t| \leq 1/N$, where $2^{k-1}/N \leq |t| \leq 2^k/N$ for $1 \leq k \leq \log_2(N^{3/4})$, and where $1/N^{1/4} \leq |t| \leq \pi$. Using the better of the bounds in (a) in each piece, prove the statement that (b) says needs to be shown.

Problems 8–12 concern singular Stieltjes measures, which for notational convenience we assume are continuous singular. In all these problems it is assumed that μ is a continuous singular measure and m is Lebesgue measure. Among other things these problems prove that the indefinite integral of μ has derivative 0 almost everywhere with respect to Lebesgue measure, i.e., $\frac{d}{dx}\int_0^x d\mu(t) = 0$ a.e. $[dx]$, with the tools of Chapter VI and without Theorem 7.2.

8. If $\epsilon > 0$ is given, prove by considering $m + \mu$ that there exists an open set U in $\mathbb{R}^1$ such that $\mu(U) < \epsilon$ and $m(U^c) = 0$.
9. If U is an open subset of $\mathbb{R}^1$ and ν is a Stieltjes measure with $\nu(U) = 0$, prove that $\lim_{h\downarrow 0}\,(2h)^{-1}\nu((x-h, x+h)) = 0$ for all x in U.
10. Let ν be any finite Stieltjes measure, and define
$$\nu^*(x) = \sup_{h>0}\,(2h)^{-1}\nu((x-h, x+h)).$$
Prove for each $\xi > 0$ that $m\big\{x \bigm| \nu^*(x) > \xi\big\} \le 5\nu(\mathbb{R}^1)/\xi$ by imitating the proof of Theorem 6.38.
11. For the singular measure μ, assume that $\mu(\mathbb{R}^1)$ is finite. Let $\epsilon > 0$ be given, and choose an open set U as in Problem 8. Define Stieltjes measures μ_1 and μ_2 by $\mu_1(A) = \mu(A \cap U)$ and $\mu_2(A) = \mu(A - U)$. Use Problem 9 to prove that $\lim_{h\downarrow 0}\,(2h)^{-1}\mu_2((x-h, x+h)) = 0$ a.e. $[dx]$, and use Problem 10 to prove for all $\xi > 0$ that
$$m\big\{x \bigm| \limsup_{h\downarrow 0}\,(2h)^{-1}\mu_1((x-h, x+h)) > \xi\big\} \le 5\epsilon/\xi.$$
12. Deduce from Problem 11 that $\lim_{h\downarrow 0}\,(2h)^{-1}\mu((x-h, x+h)) = 0$ a.e. $[dx]$. By reviewing the proof of Corollary 6.40, show how the argument in Problems 8–11 can be adjusted to yield the better conclusion that $\frac{d}{dx}\int_0^x d\mu(t) = 0$ a.e. $[dx]$.

CHAPTER VIII

Fourier Transform in Euclidean Space

Abstract. This chapter develops some of the theory of the $\mathbb{R}^N$ Fourier transform as an operator that carries certain spaces of complex-valued functions on $\mathbb{R}^N$ to other spaces of such functions.

Sections 1–3 give the indispensable parts of the theory, beginning in Section 1 with the definition, the fact that integrable functions are mapped to bounded continuous functions, and various transformation rules. In Section 2 the main results concern L^1, chiefly the vanishing of the Fourier transforms of integrable functions at infinity, the fact that the Fourier transform is one-one, and the all-important Fourier inversion formula. The third section builds on these results to establish a theory for L^2. The Fourier transform carries functions in $L^1 \cap L^2$ to functions in L^2, preserving the L^2 norm; this is the Plancherel formula. The Fourier transform therefore extends by continuity to all of L^2, and the Riesz–Fischer Theorem says that this extended mapping is onto L^2. These results allow one to construct bounded linear operators on L^2 commuting with translations by multiplying by L^∞ functions on the Fourier transform side and then using Fourier inversion; a converse theorem is proved in the next section.

Section 4 discusses the Fourier transform on the Schwartz space, the subspace of L^1 consisting of smooth functions with the property that the product of any iterated partial derivative of the function with any polynomial is bounded. The Fourier transform carries the Schwartz space in one-one fashion onto itself, and this fact leads to the proof of the converse theorem mentioned above.

Section 5 applies the Schwartz space in $\mathbb{R}^1$ to obtain the Poisson Summation Formula, which relates Fourier series and the Fourier transform. A particular instance of this formula allows one to prove the functional equation of the Riemann zeta function.

Section 6 develops the Poisson integral formula, which transforms functions on $\mathbb{R}^N$ into harmonic functions on a half space in $\mathbb{R}^{N+1}$. A function on $\mathbb{R}^N$ can be recovered as boundary values of its Poisson integral in various ways.

Section 7 specializes the theory of the previous section to $\mathbb{R}^1$, where one can associate a "conjugate" harmonic function to any harmonic function in the upper half plane. There is an associated conjugate Poisson kernel that maps a boundary function to a harmonic function conjugate to the Poisson integral. The boundary values of the harmonic function and its conjugate are related by the Hilbert transform, which implements a "90° phase shift" on functions. The Hilbert transform is a bounded linear operator on L^2 and is of weak type (1, 1).

1. Elementary Properties

Although the Fourier transform in the one-variable case dates from the early nineteenth century, it was not until the introduction of the Lebesgue integral early in the twentieth century that the theory could advance very far. Fourier series in one variable have a standard physical interpretation as representing a resolution into component frequencies of a periodic signal that is given as a function of time. In the presence of the Riesz–Fischer Theorem, they are especially handy at analyzing time-independent operators on signals, such as those given by filters. An operator of this kind takes a function f with Fourier series $f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$ into the expression $\sum_{n=-\infty}^{\infty} m_n c_n e^{inx}$, where the constants m_n depend only on the filter. If the original function f is in L^2 and if the constants m_n are bounded, the Riesz–Fischer Theorem allows one to interpret the new series as the Fourier series of a new L^2 function $T(f)$, and thus the effect of the filter is to carry f to $T(f)$.

If one imagines that the period is allowed to increase without limit, one can hope to obtain convergence of some sort to a transform that handles aperiodic signals, and this was once a common attitude about how to view the Fourier transform. In the twentieth century the Fourier transform began to be developed as an object in its own right, and soon the theory was extended from one variable to several variables.

The Fourier transform in Euclidean space $\mathbb{R}^N$ is a mapping of suitable kinds of functions on $\mathbb{R}^N$ to other functions on $\mathbb{R}^N$. The functions will in all cases now be assumed to be complex valued. The underlying $\mathbb{R}^N$ is usually regarded as space, rather than time, and the Fourier transform is of great importance in studying operators that commute with translations, i.e., spatially homogeneous operators. One example of such an operator is a linear partial differential operator with constant coefficients, and another is convolution with a fixed function. In the latter case if $\mathcal{F}$ denotes the Fourier transform and h is a fixed function, the relevant formula is $\mathcal{F}(h * f) = \mathcal{F}(h)\mathcal{F}(f)$, the product on the right side being the pointwise product of two functions. Thus convolution can be understood in terms of the simpler operation of pointwise multiplication if we understand what $\mathcal{F}$ does and we understand how to invert $\mathcal{F}$.

In the actual definition of the Fourier transform, factors of 2π invariably pop up here and there, and there is no universally accepted place to put these factors. This ambiguity is not unlike the distinction between radians and cycles in connection with frequencies in physics; again the distinction is a factor of 2π. The definition that we shall use occurs quite commonly these days, namely

$$\boxed{\widehat{f}(y) = \mathcal{F}f(y) = \int_{\mathbb{R}^N} f(x)e^{-2\pi i x\cdot y}\,dx,}$$

with $x \cdot y$ referring to the dot product and with the 2π in the exponent. The formula for $\mathcal{F}^{-1}$ will turn out to be similar looking, except that the minus sign is changed to plus in the exponent. Some authors drop the 2π from the exponent, and then a factor of $(2\pi)^{-N}$ is needed in the inversion formula. Other authors who drop the 2π from the exponent also include a factor of $(2\pi)^{-N}$ in front of the integral; then the inversion formula requires no such factor. Still other authors who drop the 2π from the exponent insert a factor of $(2\pi)^{-N/2}$ in the formula for both $\mathcal{F}$ and its inverse. In all cases, there is a certain utility in adjusting the definition of convolution by an appropriate power of 2π so that the Fourier transform of a convolution is indeed the pointwise product of the Fourier transforms. The relationships among these alternative formulas are examined in Problem 1 at the end of the chapter.

At any rate, in this book we take the boxed formula above as the definition of the Fourier transform of a function f in $L^1(\mathbb{R}^N, dx)$. Convolution was defined in Section VI.2. Although there are many elementary functions for which one can compute the Fourier transform explicitly, there are precious few for which one can make the pair of calculations that compute the Fourier transform and verify the inversion formula. One example is $e^{-\pi|x|^2}$, which will be examined in the next section.

Recall from Section VI.1 that the translate $\tau_{x_0} f$ is defined by $\tau_{x_0} f(x) = f(x - x_0)$.

Proposition 8.1. The Fourier transform on $L^1(\mathbb{R}^N)$ has these properties:

(a) f in L^1 implies that $\widehat{f}$ is bounded and uniformly continuous with $\|\widehat{f}\|_{\sup} \le \|f\|_1$,

(b) f in L^1 implies that the translate $\tau_{x_0} f$ and the product $f(x)e^{-2\pi i x\cdot y_0}$ have $(\tau_{x_0} f)\widehat{\ }(y) = e^{-2\pi i x_0\cdot y}\widehat{f}(y)$ and $\mathcal{F}(f(x)e^{2\pi i x\cdot y_0})(y) = (\tau_{y_0}\widehat{f})(y)$,

(c) f and g in L^1 implies $\widehat{f * g} = \widehat{f}\,\widehat{g}$,

(d) f in L^1 implies $\widehat{f^*} = \overline{\widehat{f}}$, where $f^*(x) = \overline{f(-x)}$,

(e) **(multiplication formula)** f and φ in L^1 implies $\int_{\mathbb{R}^N} f\widehat{\varphi}\,dx = \int_{\mathbb{R}^N}\widehat{f}\,\varphi\,dx$,

(f) f in L^1 and $2\pi i x_j f$ in L^1 implies that $\frac{\partial \widehat{f}}{\partial y_j}$ exists in the ordinary sense everywhere and satisfies $\frac{\partial \widehat{f}}{\partial y_j} = \mathcal{F}(-2\pi i x_j f)$,

(g) f in L^1 and $\frac{\partial f}{\partial x_j}$ existing in the L^1 sense, i.e., $\lim_{h\to 0} h^{-1}(\tau_{-he_j} f - f)$ existing in L^1, implies $\mathcal{F}\big(\frac{\partial f}{\partial x_j}\big)(y) = 2\pi i y_j \widehat{f}(y)$. This formula holds also when f is in $L^1 \cap C^1$, the ordinary $\frac{\partial f}{\partial x_j}$ is in L^1, and f vanishes at infinity.

PROOF. All the integrals will be over $\mathbb{R}^N$, and we drop $\mathbb{R}^N$ from the notation. For (a), we have $|\widehat{f}(y)| \le \int |f(x)|\,dx = \|f\|_1$, and hence $\|\widehat{f}\|_{\sup} \le \|f\|_1$. Also,

$$\begin{aligned} |\widehat{f}(y_1) - \widehat{f}(y_2)| &\le \int |f(x)|\,|e^{-2\pi i x\cdot y_1} - e^{-2\pi i x\cdot y_2}|\,dx \\ &= \int |f(x)|\,|e^{-2\pi i x\cdot(y_1 - y_2)} - 1|\,dx. \end{aligned}$$

On the right side the second factor of the integrand is bounded by 2 and tends to 0 for each x as $y_1 - y_2$ tends to 0. Thus the right side tends to 0 by dominated convergence at a rate depending only on $y_1 - y_2$.

For (b), $(\tau_{x_0} f)\widehat{\ }(y) = \int f(x - x_0) e^{-2\pi i x\cdot y}\,dx = \int f(x) e^{-2\pi i (x+x_0)\cdot y}\,dx = e^{-2\pi i x_0\cdot y}\widehat{f}(y)$ and

$$\begin{aligned}\mathcal{F}(f(x)e^{2\pi i x\cdot y_0})(y) &= \int f(x) e^{2\pi i x\cdot y_0} e^{-2\pi i x\cdot y}\,dx \\ &= \int f(x) e^{-2\pi i x\cdot(y-y_0)}\,dx = (\tau_{y_0}\widehat{f})(y).\end{aligned}$$

For (c), we use Fubini's Theorem. The standard technique for verifying the theorem's applicability was mentioned near the end of Section V.7. Let us see the technique in context this once. The procedure is to write out the computation, blindly making the interchange, and then to check the validity of the interchange by imagining that absolute value signs have been put in place. What needs to be verified is that the double or iterated integrals with the absolute value signs in place are finite. The computation here is

$$\begin{aligned}\widehat{f * g}(y) &= \iint f(x-t) g(t) e^{-2\pi i x\cdot y}\,dt\,dx = \iint f(x-t) g(t) e^{-2\pi i x\cdot y}\,dx\,dt \\ &= \iint f(x) g(t) e^{-2\pi i (x+t)\cdot y}\,dx\,dt = \widehat{f}(y)\widehat{g}(y).\end{aligned}$$

The steps with absolute value signs in place around the integrands are

$$\begin{aligned}\iint |f(x-t) g(t) e^{-2\pi i x\cdot y}|\,dt\,dx &= \iint |f(x-t) g(t) e^{-2\pi i x\cdot y}|\,dx\,dt \\ &= \iint |f(x) g(t) e^{-2\pi i (x+t)\cdot y}|\,dx\,dt.\end{aligned}$$

The first interchange is valid, but the first and second integrals are not so clearly finite. What is clear, because f and g are integrable, is that we have finiteness for the third integral, and the second and third integrals are equal by a translation in the inner integration. Thus the computation of $\widehat{f * g}(y)$ is justified.

For (d), we have $\widehat{f^*}(y) = \int \overline{f(-x)} e^{-2\pi i x\cdot y}\,dx = \overline{\int f(x) e^{-2\pi i x\cdot y}\,dx} = \overline{\widehat{f}(y)}$.

For (e), we use Fubini's Theorem, justifying the details in the same way as in (c). We obtain

$$\begin{aligned}\int f\widehat{\varphi}\,dx &= \iint f(x)\varphi(y) e^{-2\pi i y\cdot x}\,dy\,dx \\ &= \iint f(x)\varphi(y) e^{-2\pi i y\cdot x}\,dx\,dy = \int \widehat{f}\varphi\,dy,\end{aligned}$$

and the interchange is valid because f and φ have been assumed integrable.

For (f), we apply (b) and obtain

$$h^{-1}\big(\widehat{f}(y + he_j) - \widehat{f}(y)\big) = \mathcal{F}\big(f(x)\, h^{-1}(e^{-2\pi i h e_j\cdot x} - 1)\big)(y).$$

Application of the Mean Value Theorem to the real and imaginary parts of $h^{-1}(e^{-2\pi i h e_j \cdot x} - 1)$ shows for $|h| \le 1$ that

$$\big|\operatorname{Re}(h^{-1}(e^{-2\pi i h e_j \cdot x} - 1))\big| = |h^{-1}(1 - \cos 2\pi h x_j)| \le 2\pi |x_j|$$

and
$$\big|\operatorname{Im}(h^{-1}(e^{-2\pi i h e_j \cdot x} - 1))\big| = |h^{-1} \sin 2\pi h x_j| \le 2\pi |x_j|,$$

hence that
$$\big|h^{-1}(e^{-2\pi i h e_j \cdot x} - 1)| \le 4\pi |x_j|.$$

Since $x_j f(x)$ is assumed integrable, we have dominated convergence in the computation of the limit of $\mathcal{F}\big(f(x)\ h^{-1}(e^{-2\pi i h e_j \cdot x} - 1)\big)(y)$ as h tends to 0, and we get $\mathcal{F}\big(-2\pi i x_j f\big)(y) = \dfrac{\partial \widehat{f}}{\partial y_j}(y)$.

For the first part of (g), the assumptions and (a) give

$$\big|\mathcal{F}(h^{-1}(\tau_{-he_j} f - f))(y) - \mathcal{F}\big(\tfrac{\partial f}{\partial x_j}\big)(y)\big| \le \big\|h^{-1}(\tau_{-he_j} f - f) - \tfrac{\partial f}{\partial x_j}\big\|_1 \to 0.$$

The left side equals $\big|\widehat{f}(y)(h^{-1}(e^{2\pi i h e_j \cdot y} - 1)) - \mathcal{F}\big(\frac{\partial f}{\partial x_j}\big)(y)\big|$ by (b), and this tends to $\big|\widehat{f}(y) 2\pi i y_j - \mathcal{F}\big(\frac{\partial f}{\partial x_j}\big)(y)\big|$. Hence $\mathcal{F}\big(\frac{\partial f}{\partial x_j}\big)(y) = 2\pi i y_j \widehat{f}(y)$.

For the second part of (g), let x_j' denote the tuple of the $N-1$ variables other than x_j. Then integration by parts in the variable x_j gives

$$\begin{aligned}
\mathcal{F}\big(\tfrac{\partial f}{\partial x_j}\big)(y) &= \textstyle\int_{\mathbb{R}^{N-1}} \int_{-\infty}^{\infty} \frac{\partial f}{\partial x_j}(x) e^{-2\pi i x \cdot y}\, dx_j\, dx_j' \\
&= \textstyle\int_{\mathbb{R}^{N-1}} \lim_n \int_{-n}^{n} \frac{\partial f}{\partial x_j}(x) e^{-2\pi i x \cdot y}\, dx_j\, dx_j' \\
&= \textstyle\int_{\mathbb{R}^{N-1}} \lim_n \big[f(x) e^{-2\pi i x \cdot y}\big]_{x_j=-n}^{n}\, dx_j' \\
&\quad - \textstyle\int_{\mathbb{R}^{N-1}} \lim_n \int_{-n}^{n} f(x)(-2\pi i y_j) e^{-2\pi i x \cdot y}\, dx_j\, dx_j' \\
&= 0 + 2\pi i y_j \widehat{f}(y),
\end{aligned}$$

as asserted.

2. Fourier Transform on L^1, Inversion Formula

The main theorem of this section is the Fourier inversion formula for $L^1(\mathbb{R}^N)$. The Fourier transform for $\mathbb{R}^1$ is the analog for the line of the mapping that carries a function f on the circle to its doubly infinite sequence $\{c_k\}$ of Fourier coefficients. The inversion problem for the circle amounts to recovering f from the c_k's. We know that the procedure is to form the partial sums $s_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}$ and to look for a sense in which $\{s_n\}$ converges to f. There is no problem for the case that f is itself a trigonometric polynomial; then s_n will be equal to f for large enough n, and no passage to the limit is necessary.

The situation with the Fourier transform is different. There is no readily available nonzero integrable function on the line analogous to an exponential on the circle for which we know an inversion formula with all constants in place. In order to obtain such an inversion formula for the Fourier transform on L^1, it is necessary to be able to invert the Fourier transform of some particular nonzero function explicitly. This step is carried out in Proposition 8.2 below, and then we can address the inversion problem of $L^1(\mathbb{R}^N)$ in general. The analog for the circle of what we shall prove for the line is a rather modest result: It would say that if $\sum |c_k|$ is finite, then the sequence of partial sums converges uniformly to a function that equals f almost everywhere. The uniform convergence is a relatively trivial conclusion, being an immediate consequence of the Weierstrass M test; but the conclusion that we recover f lies deeper and incorporates a version of the uniqueness theorem.

Proposition 8.2. $\mathcal{F}\big(e^{-\pi|x|^2}\big) = e^{-\pi|y|^2}$.

REMARKS. Readers who know about the Cauchy Integral Theorem from complex-variable theory or else Green's Theorem in the theory of line integrals will recognize that the calculation below amounts to an application of one or the other of these theorems to the function $e^{-\pi z^2}$ over a long thin geometric rectangle next to the x axis in $\mathbb{C}$. However, the present application of either of these theorems is so simple that we can without difficulty substitute a proof of one of these theorems in the special case of interest, and hence neither of these other theorems needs to be assumed. As the proof below will show, matters come down to the Fundamental Theorem of Calculus in its traditional form (Theorem 1.32).

PROOF. The question is whether

$$\int_{\mathbb{R}^N} e^{-\pi(x_1^2+\cdots+x_N^2)} e^{-2\pi i(x_1y_1+\cdots+x_Ny_N)}\,dx_1\cdots dx_N \stackrel{?}{=} e^{-\pi(y_1^2+\cdots+y_N^2)},$$

and the integral on the left is the product of N integrals in one variable. Thus the question is whether

$$\int_{-\infty}^{\infty} e^{-\pi(x^2+2ixy)}\,dx \stackrel{?}{=} e^{-\pi y^2}.$$

We start by observing that

$$\int_{-\infty}^{\infty} e^{-\pi(x^2+2ixy)}\,dx = e^{-\pi y^2}\int_{-\infty}^{\infty} e^{-\pi(x+iy)^2}\,dx. \tag{$*$}$$

Write

$$e^{-\pi(x+iy)^2} = u(x,y) + iv(x,y) = e^{-\pi(x^2-y^2)}\cos 2\pi xy - ie^{-\pi(x^2-y^2)}\sin 2\pi xy.$$

Direct calculation gives[1]

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{$**$}$$

Regard n as positive and large. Then

$$\begin{aligned}
&\textstyle\int_{-n}^{n} u(s,0)\,ds - \int_{-n}^{n} u(s,y)\,ds \\
&\qquad = -\textstyle\int_{-n}^{n}\int_0^y \frac{\partial u}{\partial y}(s,t)\,dt\,ds && \text{by Theorem 1.32} \\
&\qquad = +\textstyle\int_{-n}^{n}\int_0^y \frac{\partial v}{\partial x}(s,t)\,dt\,ds && \text{by } (**) \\
&\qquad = \textstyle\int_0^y\int_{-n}^{n} \frac{\partial v}{\partial x}(s,t)\,ds\,dt && \text{by Fubini's Theorem} \\
&\qquad = \textstyle\int_0^y v(n,t)\,dt - \int_0^y v(-n,t)\,dt && \text{by Theorem 1.32.}
\end{aligned}$$

With y fixed we let n tend to infinity. Then $v(n,t)$ and $v(-n,t)$ tend to 0 uniformly for t between 0 and y by inspection of v, and hence the right side of our expression tends to 0. Thus

$$\textstyle\int_{-\infty}^{\infty} u(s,0)\,ds = \int_{-\infty}^{\infty} u(s,y)\,ds,$$

which says that

$$\operatorname{Re}\int_{-\infty}^{\infty} e^{-\pi x^2}\,dx = \operatorname{Re}\int_{-\infty}^{\infty} e^{-\pi(x+iy)^2}\,dx. \tag{$\dagger$}$$

Similarly we calculate

$$\begin{aligned}
\textstyle\int_{-n}^{n} v(s,0)\,ds - \int_{-n}^{n} v(s,y)\,ds &= -\textstyle\int_{-n}^{n}\int_0^y \frac{\partial v}{\partial y}(s,t)\,dt\,ds \\
&= -\textstyle\int_{-n}^{n}\int_0^y \frac{\partial u}{\partial x}(s,t)\,dt\,ds && \text{by } (**) \\
&= -\textstyle\int_0^y\int_{-n}^{n} \frac{\partial u}{\partial x}(s,t)\,ds\,dt \\
&= -\textstyle\int_0^y u(n,t)\,dt + \int_0^y u(-n,t)\,dt.
\end{aligned}$$

Again we can see that the right side tends to 0, and thus

$$\textstyle\int_{-\infty}^{\infty} v(s,0)\,ds = \int_{-\infty}^{\infty} v(s,y)\,ds,$$

which says that

$$\operatorname{Im}\int_{-\infty}^{\infty} e^{-\pi x^2}\,dx = \operatorname{Im}\int_{-\infty}^{\infty} e^{-\pi(x+iy)^2}\,dx. \tag{$\dagger\dagger$}$$

Taking $(*)$ into account and combining $(\dagger)$ and $(\dagger\dagger)$, we obtain

$$\int_{-\infty}^{\infty} e^{-\pi(x^2+2ixy)}\,dx = e^{-\pi y^2}\int_{-\infty}^{\infty} e^{-\pi(x+iy)^2}\,dx = e^{-\pi y^2}\int_{-\infty}^{\infty} e^{-\pi x^2}\,dx,$$

and the proposition follows from the formula $\int_{-\infty}^{\infty} e^{-\pi x^2}\,dx = 1$ given in Proposition 6.33.

[1]The equations $(**)$ are called the **Cauchy–Riemann equations**. They occur again in Section 7.

We shall use dilations to create an approximate identity out of $e^{-\pi|x|^2}$ in the style of Section VI.2. Put $\varphi(x) = e^{-\pi|x|^2}$ and define $\varphi_\varepsilon(x) = \varepsilon^{-N}\varphi(\varepsilon^{-1}x)$ for $\varepsilon > 0$. Whenever φ in an integrable function and φ_ε is formed in this way, we have

$$\begin{aligned}\widehat{\varphi_\varepsilon}(y) &= \textstyle\int_{\mathbb{R}^N} \varphi_\varepsilon(x)e^{-2\pi i x\cdot y}\,dx = \varepsilon^{-N}\int_{\mathbb{R}^N} \varphi(\varepsilon^{-1}x)e^{-2\pi i x\cdot y}\,dx \\ &= \textstyle\int_{\mathbb{R}^N} \varphi(x)e^{-2\pi i x\cdot \varepsilon y}\,dx = \widehat{\varphi}(\varepsilon y),\end{aligned}$$

the next-to-last equality following from the change of variables $\varepsilon^{-1}x \mapsto x$.

For the particular function $\varphi(x) = e^{-\pi|x|^2}$, this calculation shows that $\widehat{\varphi_\varepsilon}(y) = e^{-\pi\varepsilon^2|y|^2}$. In particular, $\widehat{\varphi_\varepsilon}$ is ≥ 0 and vanishes at ∞ for each fixed $\varepsilon > 0$. As ε decreases to 0, $\widehat{\varphi_\varepsilon}$ increases pointwise to the constant function 1. The constant c in Theorem 6.20 for this φ is $c = \int_{\mathbb{R}^N} \varphi(x)\,dx = 1$ by Proposition 6.33. That theorem gives various convergence results for $\varphi_\varepsilon * f$, one of which is that $\varphi_\varepsilon * f$ converges to f in L^1 if f is in L^1.

Theorem 8.3 (Riemann–Lebesgue Lemma). If f is in $L^1(\mathbb{R}^N)$, then the continuous function $\widehat{f}$ vanishes at infinity.

PROOF. The continuity of $\widehat{f}$ is by Proposition 8.1a. Put $\varphi(x) = e^{-\pi|x|^2}$ and form φ_ε. Then parts (c) and (a) of Proposition 8.1 give

$$\|\widehat{\varphi_\varepsilon}\widehat{f} - \widehat{f}\|_{\sup} = \|\widehat{\varphi_\varepsilon * f} - \widehat{f}\|_{\sup} \leq \|\varphi_\varepsilon * f - f\|_1,$$

and Theorem 6.20 shows that the right side tends to 0 as ε decreases to 0. Hence $e^{-\pi\varepsilon^2|y|^2}\widehat{f}(y)$ tends to $\widehat{f}(y)$ uniformly in y. Since $\widehat{f}$ is bounded (Proposition 8.1a), $e^{-\pi\varepsilon^2|y|^2}\widehat{f}(y)$ vanishes at infinity. The uniform limit of functions vanishing at infinity vanishes at infinity, and the theorem follows.

Theorem 8.4 (Fourier inversion formula). If f is in $L^1(\mathbb{R}^N)$ and $\widehat{f}$ is in $L^1(\mathbb{R}^N)$, then f can be redefined on a set of measure 0 so as to be continuous. After this adjustment,

$$f(x) = \int_{\mathbb{R}^N} \widehat{f}(y)e^{2\pi i x\cdot y}\,dy.$$

PROOF. By way of preliminaries, recall from Proposition 8.1e that the multiplication formula gives $\int f\,\widehat{g}\,dx = \int \widehat{f}\,g\,dx$ whenever f and g are both integrable. With ε fixed for the moment, let us apply this formula with $g(x) = e^{-\pi\varepsilon^2|x|^2}$. The remarks before Theorem 8.3 about how the Fourier transform interacts with dilations show that $\widehat{g}(y) = \varepsilon^{-N}e^{-\pi\varepsilon^{-2}|y|^2}$. In other words, if we take $\varphi(x) = e^{-\pi|x|^2}$, then

$$\int_{\mathbb{R}^N} f(x)\varphi_\varepsilon(x)\,dx = \int_{\mathbb{R}^N} \widehat{f}(y)e^{-\pi\varepsilon^2|y|^2}\,dy. \qquad (*)$$

To prove the theorem, consider first the special case that f is bounded and continuous. If we let ε decrease to 0 in $(*)$, the left side tends to $f(0)$ by Theorem 6.20c, and the right side tends to $\int_{\mathbb{R}^N} \widehat{f}(y)\,dy$ by dominated convergence since $\widehat{f}$ is assumed integrable. Thus $f(0) = \int_{\mathbb{R}^N} \widehat{f}(y)\,dy$. Applying this conclusion to the translate $\tau_{-x} f$ and using Proposition 8.1b, we obtain

$$f(x) = (\tau_{-x} f)(0) = \int (\tau_{-x} f)\widehat{\ }(y)\,dy = \int \widehat{f}(y)e^{2\pi i x\cdot y}\,dy,$$

as required.

Without the special assumption on f, we adjust the above argument a little. Using the equality $\varphi_\varepsilon(-y) = \varphi_\varepsilon(y)$, we apply $(*)$ to the translate $\tau_{-x} f$ of f to get

$$\begin{aligned}\int \widehat{f}(y)e^{2\pi i x\cdot y}e^{-\pi\varepsilon^2|y|^2}\,dy &= \int f(x+y)\varphi_\varepsilon(y)\,dy \\ &= \int f(x-y)\varphi_\varepsilon(y)\,dy = (\varphi_\varepsilon * f)(x).\end{aligned}$$

As ε decreases to 0, the left side tends pointwise to $\int \widehat{f}(y)e^{2\pi i x\cdot y}\,dy$ by dominated convergence, and the result is a continuous function of x, by a version of Proposition 8.1a. The right side tends to f in L^1 by Theorem 6.20, and hence Theorem 5.59 shows that a subsequence of $\varphi_\varepsilon * f$ tends to f almost everywhere. Thus $f(x) = \int_{\mathbb{R}^N} \widehat{f}(y)e^{2\pi i x\cdot y}\,dy$ almost everywhere, with the right side continuous.

Corollary 8.5. The Fourier transform is one-one on $L^1(\mathbb{R}^N)$.

PROOF. If f is in L^1 and $\widehat{f}$ is identically 0, then $\widehat{f}$ is in L^1, and the inversion formula (Theorem 8.4) applies. Hence f is 0 almost everywhere.

3. Fourier Transform on L^2, Plancherel Formula

We mentioned in Section 1 that the Fourier transform is of great importance in analyzing operators that commute with translations. The initial analysis of such operators is done on $L^2(\mathbb{R}^N)$, and this section describes some of how that analysis comes about. The first result is the theorem for $\mathbb{R}^N$ that is the analog of Parseval's Theorem for the circle.

Theorem 8.6 (Plancherel formula). If f is in $L^1(\mathbb{R}^N)\cap L^2(\mathbb{R}^N)$, then $\|\widehat{f}\|_2 = \|f\|_2$.

REMARKS. There is a formal computation that is almost a proof, namely

$$\begin{aligned}\int |f(x)|^2\,dx &= \int f^*(-x)f(x)\,dx = (f^* * f)(0) \\ &= \int \widehat{f^* * f}(y)\,dy = \int \widehat{f^*}(y)\widehat{f}(y)\,dy = \int |\widehat{f}(y)|^2\,dy,\end{aligned}$$

the middle equality using the Fourier inversion formula (Theorem 8.4). What is needed in order to make this computation into a proof is a verification that the Fourier inversion formula actually applies. We know that $f^* * f$ is in L^1 since f^* and f are in L^1, and we know from Proposition 6.18 that $f^* * f$ is continuous, being in $L^2 * L^2$. But it is not immediately obvious that the Fourier transform to which the inversion formula is to be applied, namely $\widehat{f^* * f} = |\widehat{f}|^2$, is in L^1. We handle this question by proving a lemma that is a little more general than necessary.

Lemma 8.7. Suppose f is in $L^1(\mathbb{R}^N)$, is bounded on $\mathbb{R}^N$, and is continuous at 0. If $\widehat{f}(y) \ge 0$ for all y, then $\widehat{f}$ is in $L^1(\mathbb{R}^N)$.

PROOF. Put $\varphi(x) = e^{-\pi|x|^2}$ and $\varphi_\varepsilon(x) = \varepsilon^{-N}\varphi(\varepsilon^{-1}x)$. Then the function $\varphi_\varepsilon * f$ is continuous by Proposition 6.18 since φ_ε is in L^∞ and f is in L^1, and

$$\lim_{\varepsilon\downarrow 0}(\varphi_\varepsilon * f)(0) = f(0)$$

by Theorem 6.20c. The function $\widehat{\varphi_\varepsilon}$ is in L^1, and $\widehat{f}$ is bounded. Hence $\widehat{\varphi_\varepsilon * f} = \widehat{\varphi_\varepsilon}\widehat{f}$ is in L^1. By the Fourier inversion formula (Theorem 8.4),

$$(\varphi_\varepsilon * f)(0) = \int_{\mathbb{R}^N} \widehat{f}(y)e^{-\pi\varepsilon^2|y|^2}\,dy.$$

Letting ε decrease to 0 and taking into account the monotone convergence, we obtain $f(0) = \int_{\mathbb{R}^N} \widehat{f}(y)\,dy$. Therefore $\widehat{f}$ is integrable.

PROOF OF THEOREM 8.6. The remarks after the statement of the theorem prove everything except that the Fourier transform $\widehat{f^* * f} = |\widehat{f}|^2$ is in L^1, and this step is carried out by Lemma 8.7.

Abstract linear operators between normed linear spaces were introduced in Section V.9, and Proposition 5.57 showed that boundedness is equivalent to uniform continuity. Let us make use of such operators now.

Theorem 8.6 allows us to extend the Fourier transform for $\mathbb{R}^N$ from $L^1 \cap L^2$ to L^2. In fact, Proposition 5.56 shows that $L^1 \cap L^2$ is dense in L^2. The conclusion of Theorem 8.6 implies that the linear operator $\mathcal{F}$ is bounded relative to the L^2 norms on domain and range, and hence it is uniformly continuous. Since the range space L^2 is complete (Theorem 5.59), Proposition 2.47 shows that $\mathcal{F}$ extends to a continuous map $\mathcal{F} : L^2 \to L^2$. This extended map, also called $\mathcal{F}$, is readily checked to be linear and then is a bounded linear operator satisfying $\|\mathcal{F}f\|_2 = \|f\|_2$ for all f in L^2.

If f is in $L^2(\mathbb{R}^N)$, we can use any approximating sequence from $L^1 \cap L^2$ to obtain a formula for $\mathcal{F}f$. One such is $f I_{B(R;0)}$, as R increases to infinity through some sequence. Thus

$$\mathcal{F}f(y) = \lim_{\substack{(\text{in } L^2 \text{ sense}) \\ R\to\infty}} \int_{|x|<R} f(x)e^{-2\pi i x\cdot y}\, dx.$$

Corollary 8.8. If f is in $L^1(\mathbb{R}^N)$ and g is in $L^2(\mathbb{R}^N)$, then $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$ and $\mathcal{F}(g^*) = \overline{\mathcal{F}(g)}$.

PROOF. Set $g_n = g I_{B(n;0)}$, so that g_n is in $L^1 \cap L^2$ for all n and $g_n \to g$ in L^2. Then $f * g_n \to f * g$ in L^2 since $\|f * g_n - f * g\|_2 = \|f * (g_n - g)\|_2 \leq \|f\|_1 \|g_n - g\|_2$. Therefore $\mathcal{F}(f)\mathcal{F}(g_n) = \mathcal{F}(f * g_n) \to \mathcal{F}(f * g)$ in L^2. Since $\mathcal{F}(f)$ is a bounded function and $\mathcal{F}(g_n) \to \mathcal{F}(g)$ in L^2, we see that $\mathcal{F}(f)\mathcal{F}(g_n) \to \mathcal{F}(f)\mathcal{F}(g)$ in L^2. Hence $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$. The identity $\mathcal{F}(g^*) = \overline{\mathcal{F}(g)}$ is proved similarly.

Theorem 8.9 (Riesz–Fischer Theorem). The Fourier transform operator $\mathcal{F}$ carries $L^2(\mathbb{R}^N)$ onto $L^2(\mathbb{R}^N)$.

PROOF. The operator $\mathcal{F}$ is built from the integral $\int_{\mathbb{R}^N} f(x)e^{-2\pi i x\cdot y}\, dx$. In a similar fashion, build an operator $\mathcal{I}$ from $\int_{\mathbb{R}^N} f(x)e^{2\pi i x\cdot y}\, dx$, or equivalently define $\mathcal{I}f(y) = \mathcal{F}f(-y)$. Then $\|\mathcal{I}f\|_2 = \|f\|_2$ for f in L^2. It is sufficient to prove that $\mathcal{F}\mathcal{I} = 1$ on L^2, since for any f in L^2, the equation $\mathcal{F}\mathcal{I} = 1$ implies that $\mathcal{I}f$ is a member of L^2 carried to L^2 by $\mathcal{F}$. Moreover, $\mathcal{F}\mathcal{I}$ is continuous, being bounded. It is therefore enough to prove that $\mathcal{F}\mathcal{I}f = f$ for f in a dense subspace of L^2. We shall do so for the dense subspace $L^1 \cap L^2$.

For a function f in $L^1 \cap L^2$ with the additional property that $\widehat{f}$ is in L^1 (and also L^2 by Theorem 8.6), Theorem 8.4 for $\mathcal{I}$ (or Theorem 8.4 applied to the function $f(-x)$) shows that $\mathcal{F}\mathcal{I}f = f$.

For a general f in $L^1 \cap L^2$, form $\varphi_\varepsilon * f$, where $\varphi(x) = e^{-\pi|x|^2}$. Then $\widehat{\varphi_\varepsilon * f} = \widehat{\varphi_\varepsilon}\widehat{f}$ is in $L^1 \cap L^2$; in fact, it is in L^2 by Proposition 6.14 and Theorem 8.6, and it is in L^1 because $\widehat{f}$ is bounded and $\widehat{\varphi_\varepsilon}$ is in L^1. By the special case just proved, $\mathcal{F}\mathcal{I}(\varphi_\varepsilon * f) = \varphi_\varepsilon * f$. Since $\mathcal{F}\mathcal{I}$ is continuous and $\varphi_\varepsilon * f \to f$ in L^2 by Theorem 6.20a, $\mathcal{F}\mathcal{I}f = f$. Thus $\mathcal{F}\mathcal{I}f = f$ on the dense subspace $L^1 \cap L^2$, and the proof is complete.

We shall be interested especially in bounded linear operators A on $L^2(\mathbb{R}^N)$ that commute with translations, i.e., that satisfy $A(\tau_x f) = \tau_x(Af)$ for all x in $\mathbb{R}^N$ and all f in L^2. Recall that the operator norm $\|A\|$ of a bounded linear operator on L^2 is the least C such that $\|Af\|_2 \leq C\|f\|_2$ for all f in L^2.

EXAMPLES.

(1) The translation τ_{x_0} is an example of a bounded linear operator on L^2 that commutes with translations; the commutativity in question follows from the commutativity of $\mathbb{R}^N$ as an additive group, and the equality $\|\tau_{x_0} f\|_2 = \|f\|_2$ shows that τ_{x_0} is bounded with $\|\tau_{x_0}\| = 1$. In terms of Fourier transforms, Proposition 8.1 shows that $(\tau_{x_0} f)\widehat{\ }(y) = e^{-2\pi i x_0 \cdot y} \widehat{f}(y)$.

(2) Another example of a bounded linear operator on L^2 that commutes with translations is the operator $Ag = f * g$ for fixed f in L^1. This commutes with translations by Proposition 6.15, and it is bounded with $\|A\| \le \|f\|_1$ by Proposition 6.14. Proposition 8.1 shows that $\widehat{Ag} = \widehat{f}\,\widehat{g}$.

(3) Let $M(y)$ be any L^∞ function on $\mathbb{R}^N$, and for f in L^2, define Af by $\widehat{Af} = M\widehat{f}$. The function $\widehat{f}$ is in L^2 by the Plancherel Theorem, $M\widehat{f}$ is in L^2 since M is essentially bounded, and $M\widehat{f}$ is the Fourier transform of some L^2 function by the Riesz–Fischer Theorem. We take this L^2 function to be Af. The brief formula is $Af = \mathcal{F}^{-1}(M\mathcal{F}f)$. From the inequalities $\|Af\|_2 = \|M\mathcal{F}f\|_2 \le \|M\|_\infty \|\mathcal{F}f\|_2 = \|M\|_\infty \|f\|_2$, we see that A is bounded with $\|A\| \le \|M\|_\infty$. The bounded linear operator A commutes with translations, since

$$\mathcal{F}(A(\tau_x f))(y) = \mathcal{F}(\mathcal{F}^{-1} M \mathcal{F} \tau_x f)(y) = M\mathcal{F}\tau_x f(y) = M(y) e^{-2\pi i x \cdot y} \mathcal{F}f(y)$$

and $$\mathcal{F}(\tau_x(Af))(y) = e^{-2\pi i x \cdot y} \mathcal{F}(Af)(y) = e^{-2\pi i x \cdot y} M(y)\mathcal{F}f(y).$$

One speaks of the function M as a **multiplier** on L^2. The previous two examples are instances of this construction. Example 1 has $M(y) = e^{-2\pi i x_0 \cdot y}$, and Example 2 has $M(y) = \widehat{f}(y)$. We shall see in Theorem 8.14 in the next section that *every* bounded linear operator A on L^2 commuting with translations arises from some such essentially bounded M and that $\|A\| = \|M\|_\infty$; for this reason a bounded linear operator on L^2 that commutes with translations is often called a "multiplier operator" or a "bounded multiplier operator" on L^2.

4. Schwartz Space

This section introduces the space $\mathcal{S}(\mathbb{R}^N)$ of Schwartz functions on $\mathbb{R}^N$. This space is a vector subspace of $L^1(\mathbb{R}^N)$, so that the Fourier transform is given on it by the usual concrete formula; $\mathcal{S}(\mathbb{R}^N)$ will turn out to be another space besides L^2 that is carried onto itself by the Fourier transform. Working with $\mathcal{S}(\mathbb{R}^N)$ provides a convenient way for using the Fourier transform and derivatives together, as becomes clearer when one studies partial differential equations.

If Q is a complex-valued polynomial on $\mathbb{R}^N$, define $Q(D)$ to be the partial differential operator with constant coefficients obtained by substituting, for each

j with $1 \le j \le N$, the operator $D_j = \frac{\partial}{\partial x_j}$ for x_j. A **Schwartz function** on $\mathbb{R}^N$ is a smooth function such that $P(x)Q(D)f$ is bounded for each pair of polynomials P and Q. An example is the function $e^{-\pi|x|^2}$, since its iterated partial derivatives are all of the form $R(x)e^{-\pi|x|^2}$ for some polynomial R. The **Schwartz space** $\mathcal{S} = \mathcal{S}(\mathbb{R}^N)$ is the set of all Schwartz functions.

The Schwartz space $\mathcal{S}$ is evidently a vector space, and it is closed under partial differentiation and under multiplication by polynomials. Closure under partial differentiation is in effect built into the definition. To see closure under multiplication by polynomials, it is enough to check closure under multiplication by each monomial x_j. This closure follows readily from the formula $Q(D)(x_j f) = Q^{\#}(D)f + x_j Q(D)f$, where $Q^{\#}$ is 0 or is a polynomial having degree strictly lower than Q has.

If f is a Schwartz function, then $P(x)Q(D)f$ is actually integrable, as well as bounded, for each pair of polynomials P and Q. In fact, $(1+|x|^2)^N P(x)Q(D)f$ is bounded, and therefore $P(x)Q(D)f$ is $\le$ a multiple of the integrable function $(1+|x|^2)^{-N}$. In particular, $\mathcal{S}$ is contained in L^1, L^2, and L^∞.

Finally the Fourier transform $\mathcal{F}$ carries $\mathcal{S}$ into itself. In fact, parts (f) and (g) of Proposition 8.1 give

$$P(x)Q(D)\widehat{f} = P(x)\mathcal{F}(Q(-2\pi i x)f) = \mathcal{F}(P((2\pi i)^{-1}D)Q(-2\pi i x)f),$$

and the right side is the Fourier transform of an L^1 function and therefore is bounded.

Proposition 8.10. The Fourier transform $\mathcal{F}$ is one-one from $\mathcal{S}(\mathbb{R}^N)$ onto $\mathcal{S}(\mathbb{R}^N)$, and the Fourier inversion formula holds on $\mathcal{S}(\mathbb{R}^N)$.

PROOF. Since $\mathcal{S} \subseteq L^1$, $\mathcal{F}$ is one-one on $\mathcal{S}$ as a consequence of Corollary 8.5. Since $\mathcal{F}(\mathcal{S}) \subseteq \mathcal{S} \subseteq L^1$, Theorem 8.4 shows that the Fourier inversion formula holds on $\mathcal{S}$. Let $(\mathcal{I}f)(x) = (\mathcal{F}f)(-x)$ for f in L^1. Then $\mathcal{I}(\mathcal{S}) \subseteq \mathcal{S}$. The Riesz–Fischer Theorem (Theorem 8.9) shows that $\mathcal{F}\mathcal{I} = 1$ on $L^1 \cap L^2$, and hence $\mathcal{F}\mathcal{I} = 1$ on $\mathcal{S}$ as well. Therefore if f is in $\mathcal{S}$, then $g = \mathcal{I}f$ is a member of $\mathcal{S}$ such that $\mathcal{F}g = f$, and we conclude that $\mathcal{F}$ carries $\mathcal{S}$ onto $\mathcal{S}$.

To make effective use of Proposition 8.10, we need to know that $\mathcal{S}(\mathbb{R}^N)$ is quite large, large enough so that we can shape functions suitably when we need them. For U open in $\mathbb{R}^N$, let $C^\infty_{\text{com}}(U)$ denote the vector space of smooth complex-valued functions on U whose support is a compact subset of U. It is apparent that $C^\infty_{\text{com}}(U)$ is closed under pointwise multiplication and that every member of $C^\infty_{\text{com}}(U)$ extends to a member of $C^\infty_{\text{com}}(\mathbb{R}^N)$ when set equal to 0 off U. But it is not apparent that $C^\infty_{\text{com}}(U)$ contains nonzero functions. We shall construct some.

Lemma 8.11. If δ_1 and δ_2 are given positive numbers with $\delta_1 < \delta_2$, then there exists ψ in $C^\infty_{\text{com}}(\mathbb{R}^N)$ such that $\psi(x)$ depends only on $|x|$, ψ is nonincreasing in $|x|$, ψ takes values in $[0, 1]$, $\psi(x) = 1$ for $|x| \le \delta_1$, and $\psi(x) = 0$ for $|x| \ge \delta_2$.

PROOF. We begin from the statement in Section I.7 that the function $f : \mathbb{R} \to \mathbb{R}$ with $f(t)$ equal to e^{-1/t^2} for $t > 0$ and equal to 0 for $t \le 0$ is smooth everywhere, including at $t = 0$. (The verification that f is smooth occurs in Problems 20–22 at the end of Chapter I.) If $\delta > 0$, then it follows that the function $g_\delta(t) = f(\delta + t) f(\delta - t)$ is smooth. Consequently the function $h_\delta(t) = \int_{-\infty}^t g_\delta(s)\, ds$ is smooth, is nondecreasing, is 0 for $t \le -\delta$, is some positive constant for $t \ge \delta$, and takes only values between 0 and that positive constant. Forming the function $h_{\delta,r}(t) = h_\delta(r + t) h_\delta(r - t)$ with r at least 3δ and dilating it suitably, we obtain a smooth even function $\psi_0(t)$ with values in $[0, c]$, the function being identically 0 for $|t| \ge \delta_2$ and being identically c for $|t| \le \delta_1$. Putting $\psi(x) = c^{-1}\psi_0(|x|)$, we obtain the desired function.

Proposition 8.12. If K and U are subsets of $\mathbb{R}^N$ with K compact, U open, and $K \subseteq U$, then there exists $\varphi \in C^\infty_{\text{com}}(U)$ with values in $[0, 1]$ such that φ is identically 1 on K.

PROOF. There is no loss of generality in assuming that K is nonempty and U is bounded. The continuous distance function $D(x, U^c)$ is everywhere positive on the compact set K and hence assumes a positive minimum c. Define K' to be the set $\{x \in \mathbb{R}^N \mid D(x, K) \le \tfrac14 c\}$. This set is compact, contains K, and has nonempty interior. Since the interior is nonempty, K' has positive Lebesgue measure $|K'|$. Applying Lemma 8.11, let h be a nonnegative smooth function that vanishes identically for $|x| \ge \tfrac14 c$ and has total integral 1.

Define $\varphi = h * I_{K'}$, where $I_{K'}$ is the indicator function of K'. Corollary 6.19 shows that φ is smooth. The function φ is ≥ 0 and has $\sup|\varphi| \le \|h\|_1 \|I_{K'}\|_\infty = 1$.

We have $\varphi(x) = \int_{\mathbb{R}^N} h(x - y) I_{K'}(y)\, dy$. If x is in K and $h(x - y)$ is nonzero, then $|x - y| \le \tfrac14 c$. Then $D(y, K) \le |x - y| \le \tfrac14 c$, and y is in K'. Hence $I_{K'}(y) = 1$, and $\varphi(x) = \int_{\mathbb{R}^N} h(x - y)\, dy = 1$.

Next, suppose $D(x, U^c) \le \tfrac14 c$ and $h(x - y)$ is nonzero, so that again $|x - y| \le \tfrac14 c$. The claim is that y is not in K', i.e., that $D(y, K) > \tfrac14 c$. Assuming the contrary, we can find, because of the compactness of K, some $k \in K$ with $|y - k| \le \tfrac14 c$. Then every $u^c \in U^c$ satisfies $c \le |u^c - k| \le |u^c - x| + |x - y| + |y - k| \le |u^c - x| + \tfrac14 c + \tfrac14 c$, and we obtain $|u^c - x| \ge \tfrac12 c$. Taking the infimum over u^c, we obtain $D(x, U^c) \ge \tfrac12 c$, and this is a contradiction. Thus y is not in K', and the integrand is identically 0 whenever $D(x, U^c) \le \tfrac14 c$. Hence $\varphi(x) = 0$ if $D(x, U^c) \le \tfrac14 c$, and the support of φ is a compact subset of U. This completes the proof.

Every function in $C^\infty_{\text{com}}(\mathbb{R}^N)$ is the Fourier transform of some Schwartz function by Proposition 8.10, and there are many such functions by Proposition 8.12. With this fact in hand, we can prove the theorem about operators commuting with translations that was promised in the previous section. We begin with a lemma.

Lemma 8.13. If A is a bounded linear operator on $L^2(\mathbb{R}^N)$ commuting with translations, then A commutes with convolution by any L^1 function.

PROOF. We are to show that $A(f*g) = f*(Ag)$ if f is in L^1 and g is in L^2. Let $\epsilon > 0$ be given. Corollary 6.17, with $g_1 = g$ and $g_2 = Ag$, shows that there exist $y_1, \dots, y_n$ in $\mathbb{R}^N$ and constants $c_1, \dots, c_n$ such that $\big\| f*g - \sum_{j=1}^n c_j \tau_{y_j} g \big\|_2 < \epsilon$ and $\big\| f*Ag - \sum_{j=1}^n c_j \tau_{y_j} Ag \big\|_2 < \epsilon$. Then we have

$$\begin{aligned} \|A(f*g) - f*Ag\|_2 \le{} & \big\| A\big(f*g - \textstyle\sum_{j=1}^n c_j\tau_{y_j} g\big)\big\|_2 \\ & + \big\| A\big(\textstyle\sum_{j=1}^n c_j \tau_{y_j} g\big) - \textstyle\sum_{j=1}^n c_j\tau_{y_j} Ag\big\|_2 \\ & + \big\| \textstyle\sum_{j=1}^n c_j\tau_{y_j} Ag - f*Ag\big\|_2. \end{aligned}$$

The first term on the right side is $\le \|A\| \, \big\| f*g - \sum_{j=1}^n c_j\tau_{y_j} g\big\| \le \epsilon\|A\|$, the second term is 0 since A commutes with translations, and the third term is $< \epsilon$ by construction.

Theorem 8.14. If A is a bounded linear operator on $L^2(\mathbb{R}^N)$ commuting with translations, then there exists an L^∞ function M such that $Af = \mathcal{F}^{-1}(M\mathcal{F}f)$ for all f in L^2. As a member of L^∞, M is unique and satisfies $\|M\|_\infty = \|A\|$.

REMARKS. The idea of the proof comes from the corresponding result for L^2 of the circle, where it is easy to define M. Call the operator T in the case of the circle. Each function e^{ikx} is in L^2, and the given operator T satisfies $\tau_{x_0}(T(e^{ikx})) = T(\tau_{x_0}(e^{ikx})) = T(e^{ik(x-x_0)}) = e^{-ikx_0}T(e^{ikx})$. If we write f for the L^2 function $T(e^{ikx})$ and form the Fourier series expansion $f(x) \sim \sum c_n e^{inx}$, then $\tau_{x_0} f$ has Fourier series $\tau_{x_0} f(x) \sim \sum c_n e^{-inx_0}e^{inx}$ by linearity and boundedness of τ_{x_0}. Since we have just seen that $\tau_{x_0} f = e^{-ikx_0} f$, we conclude that $\sum c_n e^{-inx_0}e^{inx} = \sum c_n e^{-ikx_0} e^{inx}$. If $c_n \ne 0$ for some n unequal to k, then we do not have the term-by-term match required by the uniqueness theorem. Hence only c_k can be nonzero, and we have $T(e^{ikx}) = c_k e^{ikx}$. The number c_k is the value of the multiplier M at the integer k. In the actual setting of the theorem, the circle is replaced by $\mathbb{R}^N$, and individual exponential functions are not in L^2. Thus this easy process for obtaining M is not available, and we are led to construct M by successive approximations.

PROOF. Choose, by Proposition 8.12, functions $\Phi_k \in C^\infty_{\text{com}}(\mathbb{R}^N)$ with

(i) $0 \le \Phi_k(y) \le 1$ for all y,

(ii) $\Phi_k(y) = 0$ for $|y| \geq k+1$,
(iii) $\Phi_k(y) = 1$ for $|y| \leq k$.

Then $\Phi_j \Phi_k = \Phi_{\min(j,k)}$ if $j \neq k$, and Φ_k is in $C^\infty_{\text{com}}(\mathbb{R}^N)$ and hence in the Schwartz space $\mathcal{S}(\mathbb{R}^N)$. Put $\varphi_k = \mathcal{F}^{-1}(\Phi_k)$. Proposition 8.10 shows that φ_k is in $\mathcal{S}$, and therefore φ_k is in $L^1 \cap L^2$. Since the Fourier transform carries convolution into pointwise product, we have $\varphi_j * \varphi_k = \varphi_{\min(j,k)}$ if $j \neq k$. Define

$$M_k = \mathcal{F}(A\varphi_k)$$

as an L^2 function. Lemma 8.13 shows that A commutes with convolution by an L^1 function, and thus $\varphi_k * A\varphi_{k+1} = A(\varphi_k * \varphi_{k+1}) = A\varphi_k = A(\varphi_k * \varphi_{k+2}) = \varphi_k * A\varphi_{k+2}$. Consequently

$$\Phi_k M_{k+1} = \Phi_k M_{k+2}$$

and
$$M_{k+1}(y) = M_{k+2}(y) \quad \text{for } |y| \leq k.$$

This equation shows that if we put

$$M(y) = M_{k+1}(y) \quad \text{for } |y| \leq k,$$

then M is consistently defined and is locally in L^2.

Let $\mathcal{S}_0 = \mathcal{F}^{-1}(C^\infty_{\text{com}}(\mathbb{R}^N)) \subseteq \mathcal{S}(\mathbb{R}^N)$. If a member f of $\mathcal{S}_0$ has $\widehat{f}(y) = 0$ for $|y| \geq k$, then $\widehat{f}\Phi_{k+1} = \widehat{f}$ and hence $f * \varphi_{k+1} = f$. Application of A gives $Af = A(f * \varphi_{k+1}) = f * A\varphi_{k+1}$. If we take the L^2 Fourier transform of both sides and use Corollary 8.8, we obtain $\mathcal{F}(Af) = M_{k+1}\widehat{f}$. The right side equals $M\widehat{f}$ since $\widehat{f}(y) = 0$ for $|y| \geq k$, and thus

$$\mathcal{F}(Af) = M\widehat{f}$$

whenever f is in $\mathcal{S}_0$ and $\widehat{f}(y) = 0$ for $|y| \geq k$.

The subspace $C^\infty_{\text{com}}(\mathbb{R}^N)$ of L^2 is dense by Corollary 6.19 and Theorem 6.20a. Since the L^2 Fourier transform carries L^2 onto L^2 and preserves norms (Theorems 8.6 and 8.9), $\mathcal{S}_0$ is dense in L^2. Let a general f in L^2 be given, and choose a sequence $\{f_j\}$ in $\mathcal{S}_0$ with $f_j \to f$ in L^2. Then $\mathcal{F}(Af_j) \to \mathcal{F}(Af)$ in L^2. By Theorem 5.59 we can pass to a subsequence, still written as $\{f_j\}$, so that $f_j \to f$ and $\mathcal{F}(Af_j) \to \mathcal{F}(Af)$ almost everywhere. Consequently

$$\begin{aligned}\mathcal{F}(Af)(y) &= \lim \mathcal{F}(Af_j)(y) = \lim M(y)\mathcal{F}(f_j)(y) \\ &= M(y)\lim \mathcal{F}(f_j)(y) = M(y)\mathcal{F}(f)(y)\end{aligned}$$

almost everywhere.

To see that M is in L^∞, suppose that $|M(y)| \geq C$ occurs at least on a set E of positive finite measure. Then I_E is in L^2. If we put $f = \mathcal{F}^{-1}(I_E)$, then we have $\|A\|\|f\|_2 \geq \|Af\|_2 = \|\mathcal{F}(Af)\|_2 = \|M\mathcal{F}(f)\|_2 = \|MI_E\|_2 \geq C\|I_E\|_2 = C\|f\|_2$, and hence $\|A\| \geq C$. Therefore $\|A\| \geq \|M\|_\infty$. In particular, M is in L^∞.

In the reverse direction we have $\|Af\|_2 = \|\mathcal{F}(Af)\|_2 = \|M\mathcal{F}(f)\|_2 \leq \|M\|_\infty\|\mathcal{F}(f)\|_2 = \|M\|_\infty\|f\|_2$ for all f in L^2, and thus $\|A\| \leq \|M\|_\infty$. We conclude that $\|M\|_\infty = \|A\|$. This completes the proof of existence.

If we have two candidates for the multiplier, say M and M_1, then subtraction of the equations $\mathcal{F}(Af) = M\mathcal{F}(f)$ and $\mathcal{F}(Af) = M_1\mathcal{F}(f)$ shows that $0 = (M - M_1)\mathcal{F}(f)$ for all f in L^2. Therefore $M = M_1$ almost everywhere. This proves uniqueness.

5. Poisson Summation Formula

The Poisson Summation Formula is a result combining Fourier series and the Fourier transform in a way that has remarkable applications, both pure and applied. Nowadays the formula is expressed as a result about Schwartz functions and therefore fits at this particular spot in the discussion of the Fourier transform.

Part of the power of the formula comes about because it applies to more settings than originally envisioned. The Euclidean version applies to the additive group $\mathbb{R}^N$, the discrete subgroup of points with integer coordinates, and the quotient group equal to the product of circle groups. In this section we shall take $N = 1$ simply because a theory of Fourier series has been developed in this book only in one variable.

We begin by stating and proving the 1-dimensional version of the theorem.

Theorem 8.15 (Poisson Summation Formula). If f is in the Schwartz space $\mathcal{S}(\mathbb{R}^1)$, then

$$\sum_{n=-\infty}^{\infty} f(x+n) = \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{2\pi inx}.$$

PROOF. Define $F(x) = \sum_{n=-\infty}^{\infty} f(x+n)$. From the definition of $\mathcal{S}$, it is easy to check that this series is uniformly convergent on any bounded interval and also the series of k^{th} derivatives is uniformly convergent on any bounded interval for each k. Consequently the function F is well defined and smooth, and it is periodic of period one. We form its Fourier series, taking into consideration that the period is 1 rather than 2π; the relevant formulas for Fourier series when the period is L rather than 2π are

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{2\pi inx/L} \qquad \text{with } c_n = \frac{1}{2L}\int_{-L}^{L} f(t)e^{-2\pi int/L}\,dt.$$

A smooth periodic function is the sum of its Fourier series, and thus

$$F(x) = \sum_{n=-\infty}^{\infty} \Big(\int_0^1 F(t) e^{-2\pi i n t}\, dt \Big) e^{2\pi i n x}. \tag{$*$}$$

The Fourier coefficient in parentheses in $(*)$ is

$$\begin{aligned} \int_0^1 F(t)e^{-2\pi int}\,dt &= \int_0^1 \sum_{k=-\infty}^{\infty} f(t+k)e^{-2\pi int}\,dt \\ &= \sum_{k=-\infty}^{\infty} \int_0^1 f(t+k)e^{-2\pi int}\,dt \\ &= \sum_{k=-\infty}^{\infty} \int_k^{k+1} f(t)e^{-2\pi int}\,dt \\ &= \int_{-\infty}^{\infty} f(t)e^{-2\pi int}\,dt \\ &= \widehat{f}(n), \end{aligned}$$

and the theorem follows by substituting this equality into $(*)$.

Corollary 8.16. $\sum_{n=-\infty}^{\infty} e^{-\pi r^{-2} n^2} = r \sum_{n=-\infty}^{\infty} e^{-\pi r^2 n^2}$ for any $r > 0$.

PROOF. The remarks before Theorem 8.3 show that it we define $\varphi(x) = e^{-\pi x^2}$ and $\varphi_\varepsilon(x) = \varepsilon^{-1}\varphi(\varepsilon^{-1}x)$, then $\widehat{\varphi_\varepsilon}(y) = \widehat{\varphi}(\varepsilon y)$. If we put $f(x) = r\varphi_r(x) = e^{-\pi r^{-2} x^2}$, then it follows that $\widehat{f}(y) = re^{-\pi r^2 x^2}$. Applying Theorem 8.15 to this f and setting $x = 0$ gives the asserted equality.

In one especially significant application of the 1-dimensional Euclidean version of the Poisson Summation Formula to pure mathematics, the remarkable identity in Corollary 8.16 can be combined with some complex-variable theory to obtain a functional equation for the **Riemann zeta function**, which is initially defined for complex s with $\operatorname{Re} s > 1$ by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \Big(1 - \frac{1}{p^s}\Big)^{-1}.$$

The functional equation relates $\zeta(s)$ to $\zeta(1-s)$. More precisely the function $\zeta(s)$ extends to be defined in a natural way[2] for all s in $\mathbb{C} - \{1\}$, and the functional equation is

$$\Lambda(1-s) = \Lambda(s), \qquad \text{where } \Lambda(s) = \zeta(s)\Gamma\big(\tfrac{1}{2}s\big)\pi^{-\frac{1}{2}s}.$$

This implication is just the beginning of a deep theory in which Fourier analysis, complex-variable theory, algebraic number theory, and algebraic geometry come

[2]The natural way is as an analytic function in $\mathbb{C} - \{1\}$. The function has a simple pole at $s = 1$.

together to yield a vast array of surprising results about prime numbers. The derivation of the functional equation of the Riemann zeta function uses some complex-variable theory, and we shall not give it.

In real-world applications the 1-dimensional Fourier transform is of great significance because of its interpretation in signal processing. A given function $f(t)$ on $\mathbb{R}^1$ is interpreted as the voltage of some signal, written as a function of time, and the Fourier transform $\widehat{f}(\omega)$ gives the components of the signal at each frequency ω. The Plancherel formula states the comforting fact that energy can be computed either by summing the contributions over time or by summing the contributions over frequency, and the result is the same. Convolutions are of special significance in the theory because they represent the effects of time-independent operations on the signal—such as the passage of the signal through a filter.

To make numerical computations, one takes some discretized version of the signal, obtained for example by rapid sampling over a long interval of time. The discrete signal, which may well be obtained at 2^n points for some n, is then regarded as periodic. In other words, the signal is really a function on a cyclic group of order 2^n. Computing a convolution involves multiplying each translate of one function by the other function at 2^n points, adding, and assembling the results. The number of steps is on the order of 2^{2n}. Alternatively, a convolution can be computed using Fourier transforms: One computes the Fourier transform of each function, does a pointwise multiplication of the new functions, and then computes an inverse Fourier transform. The pointwise multiplication involves only 2^n steps, which is relatively trivial compared with 2^{2n} steps. How many steps are involved in the computation of a Fourier transform? Naively it would seem that an exponential depending on y has to be multiplied by the value of the function at each point x and the results added, hence 2^{2n} steps. However, the mechanism of the Poisson Summation Formula contains a better way of carrying out the computation of the Fourier transform that involves only about $n2^n$ steps. The algorithm in question is known as the **fast Fourier transform** and is discussed in more detail in Problems 13–18 at the end of the chapter. The upshot is that the Poisson Summation Formula leads to a practical device that cuts down enormously on the cost of analyzing signals mathematically.

Although the Poisson Summation Formula as stated in this section relates the real line, the subgroup of integers, and the quotient circle group, the fast Fourier transform iterates versions of the formula for settings that are different from this. The groups in question are cyclic of order 2^k with $k \leq n$. We can take the subgroup to have order 2, and the quotient group then has order 2^{k-1}. A still more general version of the Poisson Summation Formula applies to any "locally compact" abelian group with a discrete subgroup having compact quotient. This more general version of the formula is used in the full-fledged application to pure mathematics that combines Fourier analysis, complex-variable theory, algebraic number theory, and algebraic geometry.

6. Poisson Integral Formula

Let $\mathbb{R}^{N+1}_+$ be the open half space $\{(x,t) \mid x \in \mathbb{R}^N \text{ and } t > 0\}$. We view the boundary $\{(x,0) \mid x \in \mathbb{R}^N\}$ as $\mathbb{R}^N$. For a function f in $L^p(\mathbb{R}^N)$ for p equal to 1, 2, or ∞, we consider the problem of finding $u(x,t)$ that is defined on $\mathbb{R}^{N+1}_+$, has f as boundary value in a suitable sense, and is **harmonic** in the sense of being a C^2 function satisfying the **Laplace equation** $\Delta u = 0$, where Δ is the **Laplacian**

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_N^2}.$$

We studied the corresponding problem for the unit disk in a sequence of problems at the ends of Chapters I, III, IV, and VI. In that situation the open disk played the role of the open half space, and the circle played the role of the Euclidean-space boundary. We were able to see that the unique possible answer, at least if f is of class C^2, is given by the Poisson integral formula for the unit disk:

$$u(r,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta-\varphi) P_r(\varphi)\,d\varphi,$$

where $P_r(\theta) = \frac{1-r^2}{1-2r\cos\theta+r^2} = \sum_{n=-\infty}^{\infty} r^{|n|}e^{in\theta}$.

The situation with $\mathbb{R}^{N+1}_+$ is different. One complication is that the boundary is not compact, and a discrete sum can no longer be expected. Another is that the harmonic function with given boundary values need not be unique; in fact, the function $u(x,t) = t$ is a nonzero harmonic function with boundary values given by $f = 0$, and thus we cannot expect to get a unique solution to a boundary-value problem unless we impose some further condition on u. In effect, the condition we impose will amount to a growth condition on the behavior of u at infinity. A partial compensation for these two complications is that the boundary is now the Euclidean space $\mathbb{R}^N$, and dilations are available as a tool.

Let us make a heuristic calculation to look for a harmonic function with given boundary values. Suppose $u(x,t)$ is the solution we seek that corresponds to f. Then we expect that the translate $\tau_{x_0}u(x,t)$ is the solution corresponding to $\tau_{x_0}f(x)$. We might further expect that the mapping $f \mapsto u(\,\cdot\,,t)$ is bounded on $L^2(\mathbb{R}^N)$. By Theorem 8.14 we would therefore have

$$\widehat{u}(y,t) = m_t(y)\widehat{f}(y),$$

for some multiplier $m_t(y)$; the Fourier transform is to be understood as occurring in the x variable only. If $t_1 > 0$ is fixed, then $u(x, t+t_1)$ is harmonic with boundary value $u(x, t_1)$, and so $\widehat{u}(y, t+t_1) = m_t(y)\widehat{u}(y, t_1) = m_t(y)m_{t_1}(y)\widehat{f}(y)$. The left side equals $m_{t+t_1}(y)\widehat{f}(y)$, and therefore

$$m_{t+t_1}(y) = m_t(y)m_{t_1}(y).$$

Since this is only a heuristic calculation anyway, we might as well assume that m is jointly measurable. Then we deduce that

$$m_t(y) = e^{tg(y)}$$

for some L^∞ function g. To compute g, we use the condition $\Delta u = 0$ more explicitly. Formally, as a result of the Fourier inversion formula, $u(x, t)$ is given as

$$\int_{\mathbb{R}^N} \widehat{u}(y, t)e^{2\pi i x\cdot y}\,dy = \int_{\mathbb{R}^N} m_t(y)\widehat{f}(y)e^{2\pi i x\cdot y}\,dy = \int_{\mathbb{R}^N} e^{tg(y)}\widehat{f}(y)e^{2\pi i x\cdot y}\,dy.$$

Without regard to the validity of the interchange of limits, we differentiate under the integral sign to obtain

$$\frac{\partial^2}{\partial x_j^2}u(x, t) = -4\pi^2 \int_{\mathbb{R}^N} y_j^2 e^{tg(y)}\widehat{f}(y)e^{2\pi i x\cdot y}\,dy$$

and
$$\frac{\partial^2}{\partial t^2}u(x, t) = \int_{\mathbb{R}^N} g(y)^2 e^{tg(y)}\widehat{f}(y)e^{2\pi i x\cdot y}\,dy.$$

Summing the derivatives and taking into account that $\widehat{f}(y)$ is rather arbitrary, we conclude that $g(y)^2 = 4\pi^2|y|^2$. Since $m_t(y)$ is an L^∞ function, we expect that the negative square root is to be used for all y. Thus $g(y) = -2\pi|y|$. Therefore our guess for the multiplier is

$$m_t(y) = e^{-2\pi t|y|}.$$

This is an L^1 function, and we begin our investigation of the validity of this answer by computing its "inverse Fourier transform," to see what to expect for the form of the bounded linear operator $f \mapsto u(\,\cdot\,, t)$.

Lemma 8.17. For $t > 0$,

$$\int_{\mathbb{R}^N} e^{-2\pi t|y|}e^{2\pi i x\cdot y}\,dy = \frac{c_N\, t}{(t^2 + |x|^2)^{\frac{1}{2}(N+1)}},$$

where $c_N = \pi^{-\frac{1}{2}(N+1)}\Gamma\big(\frac{N+1}{2}\big)$.

REMARK. The idea is to handle $t = 1$ first and then to derive the formula for other t's by taking dilations into account. For $t = 1$, we express $e^{-2\pi|y|}$ as an integral of dilates of $e^{-\pi|y|^2}$, and then the integral in question will be computable in terms of the known inverse Fourier transforms of dilates of $e^{-\pi|y|^2}$.

PROOF. In one dimension, direct calculation using calculus methods on the intervals $[0, +\infty)$ and $(-\infty, 0]$ separately gives

$$\int_{-\infty}^{\infty} e^{-2\pi|u|} e^{-2\pi i u v}\, du = \frac{1}{\pi}\,\frac{1}{1+v^2}.$$

Since $(1+v^2)^{-1}$ is integrable, the Fourier inversion formula in $\mathbb{R}^1$ (Theorem 8.4) then shows that

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{1+v^2}\, e^{2\pi i u v}\, dv = e^{-2\pi|u|}.$$

Putting $u = |y|$ with y in $\mathbb{R}^N$ yields

$$e^{-2\pi|y|} = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{2\pi i v|y|}\,(1+v^2)^{-1}\, dv. \qquad (*)$$

Any $\beta > 0$ has $\beta^{-1} = \int_0^\infty e^{-\beta s}\, ds$, and hence $(1+v^2)^{-1} = \pi\int_0^\infty e^{-(1+v^2)\pi s}\, ds$. Substitution for $(1+v^2)^{-1}$ in $(*)$, interchange of integrals by Fubini's Theorem, and use of the formula in $\mathbb{R}^1$ for the inverse Fourier transform of a dilate of $e^{-\pi v^2}$ gives

$$e^{-2\pi|y|} = \int_0^\infty e^{-\pi s}\Big[\int_{-\infty}^{\infty} e^{2\pi i v|y|} e^{-\pi v^2 s}\, dv\Big]\, ds = \int_0^\infty e^{-\pi s} s^{-1/2} e^{-\pi|y|^2/s}\, ds,$$

and this is our formula for $e^{-2\pi|y|}$ as an integral of dilates of $e^{-\pi|y|^2}$.

We multiply both sides by $e^{2\pi i x\cdot y}$, integrate, interchange the order of integration, use the formula in $\mathbb{R}^N$ for the inverse Fourier transform of a dilate of $e^{-\pi|y|^2}$, and make a change of variables $\pi s(1+|x|^2) \to s$. The result is

$$\begin{aligned}
\int_{\mathbb{R}^N} e^{-2\pi|y|} e^{2\pi i x\cdot y}\, dy &= \int_0^\infty e^{-\pi s} s^{-1/2}\Big[\int_{\mathbb{R}^N} e^{-\pi|y|^2/s} e^{2\pi i x\cdot y}\, dy\Big]\, ds \\
&= \int_0^\infty e^{-\pi s} s^{\frac{1}{2}(N-1)} e^{-\pi s|x|^2}\, ds \\
&= \int_0^\infty e^{-\pi s(1+|x|^2)} s^{\frac{1}{2}(N-1)}\, ds \\
&= \pi^{-\frac{1}{2}(N-1)}(1+|x|^2)^{-\frac{1}{2}(N-1)}\pi^{-1}(1+|x|^2)^{-1}\int_0^\infty e^{-s} s^{\frac{1}{2}(N-1)}\, ds \\
&= \pi^{-\frac{1}{2}(N+1)}\Gamma\big(\tfrac{N+1}{2}\big)(1+|x|^2)^{-\frac{1}{2}(N+1)}.
\end{aligned}$$

The proof is completed by making use of the effect of the Fourier transform on dilations. We have just seen that the function $\varphi(x) = (1+|x|^2)^{-\frac{1}{2}(N+1)}$ is integrable with Fourier transform $c_N^{-1}\widehat{\varphi}(y) = c_N^{-1}e^{-2\pi|y|}$. Then $\varphi_t(x) = t^{-N}\varphi(t^{-1}x) = t(t^2+|x|^2)^{-\frac{1}{2}(N+1)}$ has Fourier transform $c_N^{-1}\widehat{\varphi}(ty) = c_N^{-1}e^{-2\pi t|y|}$.

We define

$$P(x,t) = P_t(x) = \frac{c_N\, t}{(t^2+|x|^2)^{\frac{1}{2}(N+1)}},$$

for $t > 0$, with c_N as in Lemma 8.17, to be the **Poisson kernel** for $\mathbb{R}^{N+1}_+$. The **Poisson integral formula** for $\mathbb{R}^{N+1}_+$ is $u(x,t) = (P_t * f)(x)$, and the function u is called the **Poisson integral** of f.

Proposition 8.18. The Poisson kernel for $\mathbb{R}^{N+1}_+$ has the following properties:

(a) $P_t(x) = t^{-n}P_1(t^{-1}x)$,
(b) P_t is integrable with $\widehat{P_t}(y) = e^{-2\pi t|y|}$,
(c) $P_t \geq 0$ and $\int_{\mathbb{R}^N} P_t(x)\,dx = 1$,
(d) $P_t * P_{t'} = P_{t+t'}$,
(e) $P(x,t)$ is harmonic in $N+1$ variables.

PROOF. Conclusion (a) is by inspection. For (b), the formula for P_t shows that P_t, for fixed t, is continuous and is of order $|y|^{-(N+1)}$ as y tends to infinity. Therefore P_t is integrable. The formula for $\widehat{P_t}$ then follows from Lemma 8.17 and the Fourier inversion formula (Theorem 8.4). In (c), the first conclusion is by inspection of the formula, and the second conclusion follows from (b) by setting $y = 0$. Conclusion (d) follows from the corresponding formula on the Fourier transform side, namely $\widehat{P_t}\widehat{P'_t} = \widehat{P_{t+t'}}$, and conclusion (e) may be verified by a routine computation.

Theorem 8.19. Let p be $1, 2$, or ∞, let f be in $L^p(\mathbb{R}^N)$, and let $u(x,t) = (P_t * f)(x)$ be the Poisson integral of f. Then

(a) u is harmonic in $N+1$ variables,
(b) $\|u(\,\cdot\,,t)\|_p \leq \|f\|_p$,
(c) $u(\,\cdot\,,t)$ converges to f in L^p as t decreases to 0 provided $p < \infty$,
(d) $u(x,t)$ converges to $f(x)$ uniformly for x in E as t decreases to 0 provided f is in L^∞ and f is uniformly continuous at the points of E,
(e) the maximal function $f^{**}(x) = \sup_{t>0}|(P_t*f)(x)|$ satisfies an inequality $m\big(\{x \mid f^{**}(x) > \xi\}\big) \leq C\|f\|_1\big/\xi$ with C independent of f and ξ,
(f) **(Fatou's Theorem)** $u(x,t)$ converges to $f(x)$ for almost every x in $\mathbb{R}^N$.

REMARKS. The theorem says that u is harmonic and has boundary value f in various senses. The hypothesis for (f) is really that f is the sum of an L^1 function and an L^∞ function, and every L^2 function has this property, as will be observed in the proof below.

PROOF. Let us leave aside (a) for the moment. Conclusion (b) is immediate from Proposition 6.14 and parts (a) and (c) of Proposition 8.18. Conclusions (c) and (d) follow from parts (a) and (c) of Theorem 6.20. Conclusion (e) follows from Corollary 6.42 and the Hardy–Littlewood Maximal Theorem (Theorem 6.38), and conclusion (f) for L^1 functions f is part of Corollary 6.42. Now suppose that f is an L^∞ function. Fix a bounded interval $[a, b]$ and write $f = f_1 + f_2$ with f_1 equal to 0 off $[a, b]$ and f_2 equal to 0 on $[a, b]$. Then $P_t * f_1$ converges almost everywhere to f_1 since f_1 is integrable, and $P_t * f_2$ converges to 0 everywhere on (a, b) by (d). Hence $P_t * f$ converges almost everywhere on (a, b); since (a, b) is arbitrary, $P_t * f$ converges almost everywhere. This proves (f).

Now we return to (a). Since $P(x, t)$ is harmonic, conclusion (a) represents an interchange of differentiation and convolution. The prototype of the tool we need is Corollary 6.19, but that result does not apply here because P_t does not have compact support. If we break a function f into two pieces, one where $|f|$ is > 1 and one where $|f|$ is ≤ 1, we see that any L^2 function is the sum of an L^1 function and an L^∞ function. Thus it is enough to prove (a) when f is in L^1 or L^∞.

Let φ be P or one of P's iterated partial derivatives of some order, let $1 \leq j \leq N + 1$, and define D_j to be $\partial/\partial x_j$ if $j \leq N$ or $\partial/\partial t$ if $j = N + 1$. It is sufficient to check that

$$h^{-1}\big((\varphi * f)((x, t) + he_j) - (\varphi * f)(x, t)\big) - ((D_j\varphi) * f)(x, t)$$

tends to 0 pointwise as h tends to 0. Taking Proposition 6.15 into account, we see that we are to check that

$$\Big(h^{-1}\big(\varphi((\,\cdot\,, t) + he_j) - \varphi(\,\cdot\,, t)\big) - (D_j\varphi)(\,\cdot\,, t)\Big) * f(x)$$

tends to 0 as h tends to 0. Proposition 6.18 shows that it is enough to have

$$h^{-1}\big(\varphi((x, t) + he_j) - \varphi(x, t)\big) - (D_j\varphi)(x, t)$$

tend to 0 in L^∞ of the x variable for each fixed t if f is in L^1, or in L^1 of the x variable for fixed t if f is in L^∞. The Mean Value Theorem shows that this expression is equal to

$$(D_j\varphi)((x, t) + h'e_j) - (D_j\varphi)(x, t)$$

for some h' between 0 and h, with h' depending on x and t, and a second application of the Mean Value Theorem shows that the expression is equal to

$$h'(D_j^2\varphi)((x, t) + h''e_j).$$

We are to show that this tends to 0, for fixed t, uniformly in x and in L^1 of the x variable as h tends to 0. It is enough to show for each fixed t that

$$(D_j^2\varphi)((x,t)+he_j) \tag{$*$}$$

is dominated in absolute value by a fixed bounded function of x and a fixed L^1 function of x when h satisfies $|h| \le \frac{1}{2}\min\{1,t\}$.

An easy induction on the degree shows that any d^{th}-order partial derivative of $P(x,t)$ is of the form $Q(x,t)(t^2+|x|^2)^{-\frac{1}{2}(N+1)-d}$, where $Q(x,t)$ is a homogeneous polynomial in (x,t) of degree $d+1$. Since any monomial of degree 1 is bounded by a multiple of $(t^2+|x|^2)^{1/2}$, the d^{th}-order partial derivative is bounded by a multiple of

$$(t^2+|x|^2)^{-\frac{1}{2}(N+1)-\frac{1}{2}(d-1)}. \tag{$**$}$$

Thus the desired properties of the expression $(*)$ will follow if it is shown that $(**)$ has these properties. This is a routine matter for $d \ge 1$, and the proof of (a) is complete.

7. Hilbert Transform

This section concerns the **Hilbert transform**, the bounded linear operator H on $L^2(\mathbb{R}^N)$ given by

$$\mathcal{F}(Hf)(y) = -i(\operatorname{sgn} y)(\mathcal{F}f)(y).$$

Formally this operator has the effect, for $y>0$, of mapping exponentials by

$$e^{2\pi i x\cdot y} \mapsto -ie^{2\pi i x\cdot y} \quad \text{and} \quad e^{-2\pi i x\cdot y} \mapsto ie^{-2\pi i x\cdot y},$$

and hence of mapping cosines and sines by

$$\cos(2\pi x\cdot y) \mapsto \sin(2\pi x\cdot y) \quad \text{and} \quad \sin(2\pi x\cdot y) \mapsto -\cos(2\pi x\cdot y).$$

For this reason, engineers sometimes call the Hilbert transform a "90° phase shift." The notion is of such importance that there is even a piece of hardware called a "Hilbert transformer" that takes an input signal and produces some kind of approximation to the Hilbert transform of the signal.[3]

We shall do some Fourier analysis in order to identify H more directly. To get an idea what H is, we begin by computing the effect on L^2 of composing the Hilbert transform and convolution with the Poisson kernel $P_\varepsilon(x)$. Then we examine what happens when ε decreases to 0.

[3]The delay in time that a Hilbert transformer requires in producing its output imposes a built-in theoretical limit for how good the approximation to the Hilbert transform can be. An exact result would require an infinite time delay.

Lemma 8.20. For $\varepsilon > 0$, $\displaystyle\int_{\mathbb{R}^1} (-i \operatorname{sgn} t) e^{-2\pi\varepsilon|t|} e^{2\pi i x\cdot t}\, dt = \frac{1}{\pi}\,\frac{x}{\varepsilon^2 + x^2}$.

PROOF. This result follows by direct calculation, using calculus methods on the intervals $[0, +\infty)$ and $(-\infty, 0]$ separately.

If we define $Q(x) = \dfrac{1}{\pi}\,\dfrac{x}{1+x^2}$ for x in $\mathbb{R}^1$, then $Q_\varepsilon(x) = \varepsilon^{-1} Q(\varepsilon^{-1} x) = \dfrac{1}{\pi}\,\dfrac{x}{\varepsilon^2 + x^2}$ is the function in the statement of Lemma 8.20. We define

$$Q(x, \varepsilon) = Q_\varepsilon(x) = \frac{1}{\pi}\,\frac{x}{\varepsilon^2 + x^2},$$

for $\varepsilon > 0$, to be the **conjugate Poisson kernel** on $\mathbb{R}^2_+$. The function Q_ε is *not* in $L^1(\mathbb{R}^1)$. However, it is in $L^2(\mathbb{R}^1)$, and therefore the convolution of Q_ε and any L^2 function is a well-defined bounded uniformly continuous function. For f in L^2, the function $v(x, \varepsilon) = (Q_\varepsilon * f)(x)$ is called the **conjugate Poisson integral** of f.

Proposition 8.21. The conjugate Poisson kernel for $\mathbb{R}^2_+$ has the following properties:

(a) the function $v(x, y) = Q_y(x)$ is harmonic in $\mathbb{R}^2_+$, and the pair of functions u and v with $u(x, y) = P_y(x)$ satisfies the **Cauchy–Riemann equations**

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},$$

(b) the L^2 Fourier transform $\mathcal{F}(Q_\varepsilon)(y)$ equals $-i(\operatorname{sgn} y)e^{-2\pi\varepsilon|y|}$,

(c) $Q_\varepsilon * P_{\varepsilon'} = Q_{\varepsilon+\varepsilon'}$.

REMARKS. A fundamental result of complex-variable theory is that if u and v are C^1 functions on an open subset of $\mathbb{C}$ satisfying the Cauchy–Riemann equations, then $f(z) = u(x, y) + iv(x, y)$ is an "analytic" function in the sense that in any open disk about any point in the open set, $f(z)$ equals the sum of a power series convergent in that disk. We shall not make use of this fact, but the analyticity of $u + iv$ is the starting point for a great deal of analysis that will not be treated in this book. In the special case of the Poisson kernel and the conjugate Poisson kernel, the function f is $f(z) = i/(\pi z)$.

PROOF. Part (a) is a routine calculation.

For (b), we know that Q_ε is in L^2 and has an L^2 Fourier transform $g = \mathcal{F}(Q_\varepsilon)$. The inverse Fourier transform $\mathcal{F}^{-1}$ on L^2 satisfies $\mathcal{F}^{-1}(g) = Q_\varepsilon$, and (b) will follow if we show that $\mathcal{F}^{-1}(f) = Q_\varepsilon$, where $f(t) = -i(\operatorname{sgn} t)e^{-2\pi\varepsilon|t|}$. For each integer $n > 0$, let $f_n(t)$ be $f(t)$ for $|t| \leq n$ and 0 for $|t| > n$. Then $f_n \to f$ in L^2 by dominated convergence, and hence $\mathcal{F}^{-1}(f_n) \to \mathcal{F}^{-1}(f)$ in L^2. By Theorem 5.59 a subsequence of $\mathcal{F}^{-1}(f_n)$ converges almost everywhere to $\mathcal{F}^{-1}(f)$. Since f is in L^1, Lemma 8.20 shows that $\mathcal{F}^{-1}(f_n)(t) = \int_{-n}^{n} f(t)e^{2\pi ixt}\,dt$ converges pointwise to $Q_\varepsilon(x)$, and therefore $\mathcal{F}^{-1}(f) = Q_\varepsilon$.

For (c), Corollary 8.8 shows that $\mathcal{F}(Q_\varepsilon * P_{\varepsilon'}) = \mathcal{F}(Q_\varepsilon)\mathcal{F}(P_{\varepsilon'})$. In combination with Proposition 8.18b, conclusion (b) of the present proposition gives $\mathcal{F}(Q_\varepsilon)(y)\mathcal{F}(P_{\varepsilon'})(y) = -i(\operatorname{sgn} y)e^{-2\pi(\varepsilon+\varepsilon')|y|}$ a.e., and this is $\mathcal{F}(Q_{\varepsilon+\varepsilon'})(y)$ a.e. by a second application of (b).

Theorem 8.22. Let f be in $L^2(\mathbb{R}^1)$, and let $u(x, y) = (P_y * f)(x)$ and $v(x, y) = (Q_y * f)(x)$ be the Poisson integral and conjugate Poisson integral of f. Then

(a) the function v is harmonic in $\mathbb{R}^2_+$, and the pair of functions u and v satisfies the Cauchy–Riemann equations,
(b) the function $Q_\varepsilon * f$ is in $L^2(\mathbb{R}^1)$ for every $\varepsilon > 0$, and its L^2 Fourier transform is $\mathcal{F}(Q_\varepsilon * f)(y) = -i(\operatorname{sgn} y)e^{-2\pi\varepsilon|y|}\mathcal{F}(f)(y)$,
(c) $\|Q_\varepsilon * f\|_2 = \|P_\varepsilon * f\|_2 \leq \|f\|_2$ for every $\varepsilon > 0$,
(d) $Q_\varepsilon * f \to H(f)$ in L^2 as ε decreases to 0.

PROOF. Conclusion (a) is handled just like Theorem 8.19a. In the proof of Theorem 8.19a, the integrability of P_ε did not play a role; it was the integrability of the iterated partial derivatives of P_ε (i.e., the case $d > 0$) that was important. The estimates involving such derivatives here are the same as in that case.

For (b), put $g = \mathcal{F}(Q_\varepsilon)\mathcal{F}(f)$. This is an L^2 function since $\mathcal{F}(Q_\varepsilon)$ is in L^∞ by inspection and since $\mathcal{F}(f)$ is in L^2 by the Plancherel formula. Define $f_n = I_{B(n;0)}f$, so that each f_n is in $L^1 \cap L^2$ and also $f_n \to f$ in L^2. Since $\mathcal{F}(Q_\varepsilon)$ is in L^∞, the Plancherel formula shows that $g_n = \mathcal{F}(Q_\varepsilon)\mathcal{F}(f_n)$ is in L^2 for each n and converges to g in L^2. Since f_n is in L^1 and Q_ε is in L^2, Corollary 8.8 gives $\mathcal{F}(Q_\varepsilon * f_n) = \mathcal{F}(Q_\varepsilon)\mathcal{F}(f_n) = g_n$ for all n. Thus $Q_\varepsilon * f_n = \mathcal{F}^{-1}(g_n)$. We now let n tend to infinity. We know that $\|Q_\varepsilon * f_n - Q_\varepsilon * f\|_{\sup} \leq \|Q_\varepsilon\|_2\|f_n - f\|_2$. Since Q_ε is in L^2 and $f_n \to f$ in L^2, $Q_\varepsilon * f_n$ converges uniformly to $Q_\varepsilon * f$. On the other hand, $\mathcal{F}^{-1}(g_n)$ converges to $\mathcal{F}^{-1}(g)$ in L^2, and Theorem 5.59 shows that a subsequence converges almost everywhere. Therefore $\mathcal{F}^{-1}(g) = Q_\varepsilon * f$. Consequently $\mathcal{F}(Q_\varepsilon * f) = g = \mathcal{F}(Q_\varepsilon)\mathcal{F}(f)$, and we obtain $\mathcal{F}(Q_\varepsilon * f)(y) = -i(\operatorname{sgn} y)e^{-2\pi\varepsilon|y|}\mathcal{F}(f)(y)$.

Conclusions (c) and (d) follow by taking L^2 Fourier transforms and using (b), Proposition 8.18b, and the Plancherel Theorem. This completes the proof.

To get a more direct formula for the Hilbert transform, we introduce the functions

$$h(x) = \begin{cases} \frac{1}{\pi x} & \text{for } |x| \geq 1, \\ 0 & \text{for } |x| < 1, \end{cases}$$

and
$$h_\varepsilon(x) = \varepsilon^{-1} h(\varepsilon^{-1} x) = \begin{cases} \frac{1}{\pi x} & \text{for } |x| \geq \varepsilon, \\ 0 & \text{for } |x| < \varepsilon. \end{cases}$$

Let $\psi(x) = Q(x) - h(x)$, so that $\psi_\varepsilon(x) = \varepsilon^{-1}\psi(\varepsilon^{-1}x) = Q_\varepsilon(x) - h_\varepsilon(x)$.

Lemma 8.23. The function ψ on $\mathbb{R}^1$ is integrable, and $\int_{-\infty}^{\infty} \psi(x)\,dx = 0$.

PROOF. For $|x| < 1$, we have $\psi(x) = Q(x) = \pi^{-1}x\big/(1+x^2)$. This is a continuous odd function, and therefore it is integrable on $[-1, 1]$ with integral 0. For $|x| \geq 1$, we have $\psi(x) = \pi^{-1}\big(\frac{x}{1+x^2} - \frac{1}{x}\big) = -\pi^{-1}\big(\frac{1}{x(1+x^2)}\big)$. This is an integrable function for $|x| \geq 1$; since it is an odd function, its integral is 0.

Theorem 8.24. Let h_ε be defined as above. If f is in $L^2(\mathbb{R}^1)$, then $h_\varepsilon * f$ is in $L^2(\mathbb{R}^1)$ for every $\varepsilon > 0$, and $h_\varepsilon * f \to H(f)$ in L^2 as ε decreases to 0.

REMARKS. More concretely the limit relation in the theorem is that

$$Hf(x) = \lim_{\substack{(\text{in } L^2 \text{ sense}) \\ \varepsilon \downarrow 0}} \frac{1}{\pi} \int_{|t| \geq \varepsilon} \frac{f(x-t)}{t}\,dt.$$

The integrand on the right side is the product of two L^2 functions on the set where $|t| \geq \varepsilon$, and it is integrable by the Schwarz inequality.

PROOF. We have $h_\varepsilon * f = Q_\varepsilon * f - \psi_\varepsilon * f$. The term $Q_\varepsilon * f$ is in L^2 by Theorem 8.22b, and the term $\psi_\varepsilon * f$ is in L^2 by Lemma 8.23 and Proposition 6.14. As ε decreases to 0, $Q_\varepsilon * f$ tends to Hf in L^2 by Theorem 8.22d, and $\psi_\varepsilon * f$ tends to 0 in L^2 by Theorem 6.20a. This completes the proof.

Now that we have the concrete formula of Theorem 8.24 for the Hilbert transform on L^2 functions, we can ask whether the Hilbert transform is meaningful on other kinds of functions. For example, we could ask, If we have some other vector space V of functions and $V \cap L^2(\mathbb{R}^1)$ is dense in V, can we extend H to V? The answer for $V = L^1(\mathbb{R}^1)$ is unfortunately negative. In fact, if f is in $L^1 \cap L^2$, then the Fourier transform $\widehat{f}$ will be continuous and the Fourier transform of Hf will have to be $-i(\operatorname{sgn} y)\widehat{f}$. If $\widehat{f}(0)$ is nonzero, then $-i(\operatorname{sgn} y)\widehat{f}$ is not continuous and cannot be the Fourier transform of an L^1 function.

However, in Chapter IX we shall introduce L^p spaces for $1 \le p \le \infty$, thereby extending the definitions we have already made for p equal to $1, 2$, and ∞. Toward the end of the chapter, we shall see that the Hilbert transform makes sense as a bounded linear operator on $L^p(\mathbb{R}^1)$ for $1 < p < \infty$. This boundedness is an indication that the Hilbert transform is not a completely wild transformation, and the result in question will be used in the problems at the end of Chapter IX to prove that the partial sums of the Fourier series of an L^p function on the circle converge to the function in L^p as long as $1 < p < \infty$.

Actually, this boundedness on L^p will be a consequence of a substitute result about L^1 that we shall prove now. Although the Hilbert transform is not a bounded linear operator on L^1, its approximations in the statement of Theorem 8.24 are of weak type $(1, 1)$, in the same sense that the passage from a function to its Hardy–Littlewood maximal function in Chapter VI was of weak type $(1, 1)$.

Theorem 8.25. Let h_1 be the function on $\mathbb{R}^1$ equal to $1/(\pi x)$ for $|x| \ge 1$ and equal to 0 for $|x| < 1$. For f in $L^1(\mathbb{R}^1) + L^2(\mathbb{R}^1)$, define

$$H_1 f(x) = h_1 * f(x) = \frac{1}{\pi} \int_{|t| \ge 1} \frac{f(x-t)}{t}\, dt$$

as the convolution of the fixed function h_1 in L^2 with a function f that is the sum of an L^1 function and an L^2 function. Then

$$\|H_1 f\|_2 \le A \|f\|_2,$$

with the constant A independent of f, and

$$m\big\{x \in \mathbb{R}^1 \;\big|\; |H_1 f(x)| > \xi\big\} \le \frac{C\|f\|_1}{\xi}$$

for every $\xi > 0$, with the constant C independent of ξ and f.

REMARK. This result about the approximation H_1 to H on L^1 and L^2 will be enough for now. The result for L^1 is much more difficult than the result for L^2. In the next chapter we shall derive from Theorem 8.25 a boundedness theorem for all the other approximations $H_\varepsilon = h_\varepsilon * (\,\cdot\,)$ on $L^p(\mathbb{R}^1)$ for $1 < p < \infty$, with a bound independent of ε, and then it will be an easy matter to get the boundedness of the Hilbert transform H itself on L^p for these values of p.

PROOF. A preliminary fact is needed that involves a computation with the function h_1. We need to know that

$$\textstyle\int_{|x| \ge 2r} |h_1(x + r') - h_1(x)|\, dx \le 6 \qquad (*)$$

whenever $0 < |r'| \le r$. To see this, we break the region of integration into four sets—one where $|x| \ge 2r$, $|x| \ge 1$, and $|x + r'| \ge 1$; a second where $|x| \ge 2r$, $|x| < 1$, and $|x + r'| \ge 1$; a third where $|x| \ge 2r$, $|x| \ge 1$, and $|x + r'| < 1$; and a fourth where $|x| \ge 2r$, $|x| < 1$, and $|x+r'| < 1$. For the fourth piece the integrand is 0. For the second and third pieces, the integrand is ≤ 1 in absolute value, and the set has measure ≤ 2; hence each of these pieces contributes at most 2. For the first piece the absolute value of the integrand is $|r'|\big/|x(x + r')| \le 2r/x^2$; thus the absolute value of the integral is $\le \int_{|x|\ge 2r} 2r/x^2\,dx = 2$. This proves $(*)$.

It is an easy matter to prove that H_1 is a bounded linear operator on L^2. In fact, $h_1 = Q_1 - \psi$, and ψ is in L^1 by Lemma 8.23. Thus Theorem 8.22c gives $\|H_1 f\|_2 \le \|Q_1 * f\|_2 + \|\psi * f\|_2 \le \|f\|_2 + \|\psi\|_1\|f\|_2$. In other words, H_1 is bounded on L^2 with $\|H_1\| \le 1 + \|\psi\|_1$. Put $A = 1 + \|\psi\|_1$.

The heart of the proof is the observation that if F is in L^1, vanishes off a bounded interval I with center y_0 and double[4] I^*, and has total integral $\int_{\mathbb{R}^1} F(y)\,dy$ equal to 0, then

$$\|H_1 F\|_{L^1(\mathbb{R}-I^*)} \le 6\|F\|_1. \qquad (**)$$

To see this, we use the fact that the total integral of F is 0 to write

$$H_1 F(x) = \int_I h_1(x - y)F(y)\,dy = \int_I [h_1(x - y) - h(x - y_0)]F(y)\,dy.$$

Taking the absolute value of both sides and integrating over $\mathbb{R} - I^*$, we obtain

$$\begin{aligned}\int_{x\notin I^*} |H_1 F(x)|\,dx &\le \int_{x\notin I^*}\int_{y\in I} |h_1(x - y) - h(x - y_0)|\,|F(y)|\,dy\,dx\\ &= \int_{y\in I}\Big[\int_{x\notin I^*} |h_1(x - y) - h(x - y_0)|\,dx\Big]|F(y)|\,dy\\ &\le 6\int_{y\in I} |F(y)|\,dy,\end{aligned}$$

the last step holding by $(*)$. This proves $(**)$.

Let the L^1 function f be given. Fix $\xi > 0$. We shall decompose the L^1 function f into the sum $f = g + b$ of a "good" function g and a "bad" function b, in a manner dependent on ξ. The good function will be in L^∞ and hence will be in $L^1 \cap L^\infty \subseteq L^2$; the effect of applying H_1 to it will be controlled by the bound of H_1 on L^2. The bad function will be nonzero on a set of rather small measure, and we shall be able to control the effect of H_1 on it by means of $(**)$.

We begin by constructing a disjoint countable system of bounded open intervals I_k such that

(i) $\sum_k m(I_k) \le 5\|f\|_1\big/\xi$,
(ii) $|f(x)| \le \xi$ almost everywhere off $\bigcup_k I_k$,
(iii) $\frac{1}{m(I_k)}\int_{I_k} |f(y)|\,dy \le 2\xi$ for each n.

[4] The "double" of a bounded interval I is an interval of twice the length of I and the same center.

Namely, let $f^*(x) = \sup_{h>0} \frac{1}{2h} \int_{[x-h,x+h]} |f(t)|\,dt$ be the Hardy–Littlewood maximal function of f, and let E be the set where $f^*(x) > \xi$. The set E is open. In fact, if $f^*(x) > \xi$, then $\frac{1}{2h} \int_{[x-h,x+h]} |f(t)|\,dt \geq \xi + \epsilon$ for some $\epsilon > 0$. For $\delta > 0$, the inequality $\frac{1}{2h} \int_{[x-h,x+h+2\delta]} |f(t)|\,dt \geq \xi + \epsilon$ shows that $f^*(x+\delta) \geq \frac{h}{h+\delta}(\xi+\epsilon)$. Hence $f^*(x+\delta) > \xi$ for δ sufficiently small. Similarly $f^*(x-\delta) > \xi$ for δ sufficiently small.

Since E is open, E is uniquely the disjoint union of countably many open intervals, and these intervals will be the sets I_k. The disjointness of the I_k's and the Hardy–Littlewood Maximal Theorem (Theorem 6.38) together give

$$\sum_k m(I_k) \leq m(E) \leq 5\|f\|_1 \big/ \xi,$$

and this proves (i) and the boundedness of the intervals. The a.e. differentiability of integrals (Corollary 6.39) shows that $|f(x)| \leq f^*(x)$ a.e., and therefore $|f(x)| \leq \xi$ a.e. off $E = \bigcup_k I_k$. This proves (ii). If $I = (a,b)$ is one of the I_k's, then a is not in E, and we must have $\frac{1}{2(b-a)} \int_{[b-2(b-a),b]} |f(t)|\,dt \leq f^*(a) \leq \xi$. Therefore $\frac{1}{b-a} \int_{[a,b]} |f(t)|\,dt \leq 2\xi$. This proves (iii).

With the open intervals I_k in hand, we define the decomposition $f = g + b$ by

$$g(x) = \begin{cases} \dfrac{1}{m(I_k)} \displaystyle\int_{I_k} f(y)\,dy & \text{for } x \in I_k, \\ f(x) & \text{for } x \notin \bigcup_k I_k, \end{cases}$$

$$b(x) = \begin{cases} f(x) - \dfrac{1}{m(I_k)} \displaystyle\int_{I_k} f(y)\,dy & \text{for } x \in I_k, \\ 0 & \text{for } x \notin \bigcup_k I_k. \end{cases}$$

Since $\big\{x \bigm| |H_1 f(x)| > \xi\big\} \subseteq \big\{x \bigm| |H_1 g(x)| > \xi/2\big\} \cup \big\{x \bigm| |H_1 b(x)| > \xi/2\big\}$, it is enough to prove

- $m\big(\big\{x \bigm| |H_1 g(x)| > \xi/2\big\}\big) \leq C'\|f\|_1 \big/ \xi$ and
- $m\big(\big\{x \bigm| |H_1 b(x)| > \xi/2\big\}\big) \leq C'\|f\|_1 \big/ \xi$

for some constant C' independent of ξ and f.

The definition of g shows that $\int_{I_k} |g(x)|\,dx \leq \int_{I_k} |f(x)|\,dx$ for all k and that $|g(x)| = |f(x)|$ for $x \notin \bigcup_k I_k$; therefore $\int_{\mathbb{R}} |g(x)\,dx \leq \int_{\mathbb{R}} |f(x)|\,dx$. Also, properties (ii) and (iii) of the I_k's show that $|g(x)| \leq 2\xi$ a.e. These two inequalities, together with the bound $\|H_1 g\|_2 \leq A\|g\|_2$, give

$$\int_{\mathbb{R}} |H_1 g(x)|^2\,dx \leq A^2 \int_{\mathbb{R}^1} |g(x)|^2\,dx \leq 2\xi A^2 \int_{\mathbb{R}} |g(x)|\,dx \leq 2\xi A^2 \int_{\mathbb{R}} |f(x)|\,dx.$$

Combining this result with Chebyshev's inequality $m\big(\{x \mid |F(x)| > \beta\}\big) \leq \beta^{-2}\int_{\mathbb{R}}|F(x)|^2\,dx$ for the function $F = H_1 g$ and the number $\beta = \xi/2$, we obtain

$$m\big(\{x \mid |H_1 g(x)| > \xi/2\}\big) \leq \frac{4}{\xi^2}\,2\xi A^2\int_{\mathbb{R}}|f(x)|\,dx = \frac{8A^2\|f\|_1}{\xi}.$$

This proves the bulleted item about g.

For b, let b_k be the product of b with the indicator function of I_k. Then we have $b = \sum_k b_k$ with the sum convergent in L^1. Since H_1 is convolution by the L^2 function h_1, $H_1 b = \sum_k H_1 b_k$ with the sum convergent in L^2. Lumping terms via Theorem 5.59 if necessary, we may assume that the convergence takes place a.e. Therefore $|H_1 b(x)| \leq \sum_k |H_1 b_k(x)|$ a.e. Using monotone convergence and $(**)$, we conclude that

$$\begin{aligned}\|H_1 b\|_{L^1\left(\mathbb{R}-\bigcup_k I_k^*\right)} &\leq \sum_k \|H_1 b_k\|_{L^1\left(\mathbb{R}-\bigcup_j I_j^*\right)}\\ &\leq \sum_k \|H_1 b_k\|_{L^1(\mathbb{R}-I_k^*)} \leq 6\|b_k\|_1 = 6\|b\|_1 \leq 6\|f\|_1.\end{aligned}$$

Thus $\quad m\big(\{x \in \mathbb{R} - \bigcup_k I_k^* \mid |H_1 b(x)| > \xi/2\}\big) \leq 6\|f\|_1\big/(\xi/2) = 12\|f\|_1/\xi.$

Since $m\big(\{x \in \bigcup_k I_k^*\}\big) \leq 5\|f\|_1\big/\xi$ by (i), we obtain $m\big(\{x \mid |H_1 b(x)| > \xi/2\}\big) \leq 17\|f\|_1\big/\xi$, and the bulleted item about b follows.

8. Problems

1. For each of the following alternative definitions of the Fourier transform in $\mathbb{R}^N$, find a constant α such that the Fourier inversion formula is as indicated, and find a constant β such that when convolution is defined by

$$f * g(x) = \beta\int_{\mathbb{R}^N} f(x-t)g(t)\,dt,$$

then the Fourier transform of the convolution is the product of the Fourier transforms.

(a) Fourier transform $\widehat{f}(y) = \int_{\mathbb{R}^N} f(x)e^{-ix\cdot y}\,dy$ and inverse Fourier transform $f(x) = \alpha\int_{\mathbb{R}^N}\widehat{f}(y)e^{ix\cdot y}\,dy$.

(b) Fourier transform $\widehat{f}(y) = (2\pi)^{-N}\int_{\mathbb{R}^N} f(x)e^{-ix\cdot y}\,dy$ and inverse Fourier transform $f(x) = \alpha\int_{\mathbb{R}^N}\widehat{f}(y)e^{ix\cdot y}\,dy$.

(c) Fourier transform $\widehat{f}(y) = (2\pi)^{-N/2}\int_{\mathbb{R}^N} f(x)e^{-ix\cdot y}\,dy$ and inverse Fourier transform $f(x) = \alpha\int_{\mathbb{R}^N}\widehat{f}(y)e^{ix\cdot y}\,dy$.

2. Let $(u, v)_2 = \int_{\mathbb{R}^N} u(x)\overline{v(x)}\,dx$ if u and v are in $L^2(\mathbb{R}^N)$, and let $\mathcal{F}$ denote the Fourier transform on $L^2(\mathbb{R}^N)$. Prove for every pair of functions f and g in L^2 that $(f, g)_2 = (\mathcal{F}(f), \mathcal{F}(g))_2$.

3. Prove that the Poisson kernel P and the conjugate Poisson kernel Q for $\mathbb{R}^2_+$ satisfy the identity $Q_\varepsilon * Q_{\varepsilon'} = P_{\varepsilon+\varepsilon'}$.

4. This problem is an analog for the Fourier transform of Problem 20c of Chapter VI concerning Fourier series and weak-star convergence. Weak-star convergence was defined in Section V.9.
   (a) If f is in $L^\infty(\mathbb{R}^N)$ and P_t is the Poisson kernel, prove that $P_t * f$ converges to f weak-star against $L^1(\mathbb{R}^N)$ as t decreases to 0. In other words, prove that $\lim_{t\downarrow 0} \int_{\mathbb{R}^N} (P_t * f)(x)g(x)\,dx = \int_{\mathbb{R}^N} f(x)g(x)\,dx$ for every g in L^1.
   (b) Theorem 8.19b shows that $\|P_t * f\|_\infty \le \|f\|_\infty$ if f is in $L^\infty(\mathbb{R}^N)$. Prove that $\lim_{t\downarrow 0} \|P_t * f\|_\infty = \|f\|_\infty$.

5. Let $\mathcal{M}^+(\mathbb{R}^N)$ be the space of finite Borel measures on $\mathbb{R}^N$. This problem introduces convolution and the Poisson integral formula for $\mathcal{M}^+(\mathbb{R}^N)$. Each finite Borel measure on $\mathbb{R}^N$ defines, by means of integration, a bounded linear functional on the normed linear space $C_{\text{com}}(\mathbb{R}^N)$ equipped with the supremum norm, and thus it is meaningful to speak of weak-star convergence of such measures against $C_{\text{com}}(\mathbb{R}^N)$.
   (a) The convolution of a finite Borel measure μ on $\mathbb{R}^N$ with an integrable function f is defined by $(f * \mu)(x) = \int_{\mathbb{R}^N} f(x - y)\,d\mu(y)$. Define the **convolution** $\mu = \mu_1 * \mu_2$ of two members of $\mathcal{M}^+(\mathbb{R}^N)$ by $\mu(E) = \int_{\mathbb{R}^N} \mu_1(E - x)\,d\mu_2(x)$ for all Borel sets E. Check that the result is a Borel measure and that the definition for $f\,dx * \mu$, i.e., for the situation in which μ_1 and μ_2 are specialized so that $\mu_1 = f\,dx$ and $\mu_2 = \mu$, is consistent with the definition in the special case.
   (b) With convolution of finite Borel measures on $\mathbb{R}^N$ defined as in (a), prove that $\int_{\mathbb{R}^N} g\,d(\mu_1 * \mu_2) = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} g(x + y)\,d\mu_1(x)\,d\mu_2(y)$ for every bounded Borel function g on $\mathbb{R}^N$.
   (c) Verify that $\|P_t * \mu\|_1 \le \mu(\mathbb{R}^N)$ if μ is in $\mathcal{M}^+(\mathbb{R}^N)$. Prove the limit formula $\lim_{t\downarrow 0} \|P_t * \mu\|_1 = \mu(\mathbb{R}^N)$.
   (d) If μ is in $\mathcal{M}^+(\mathbb{R}^N)$, prove that the measures $(P_t * \mu)(x)\,dx$ converge to μ weak-star against $C_{\text{com}}(\mathbb{R}^N)$ as t decreases to 0. In other words, prove that $\lim_{t\downarrow 0} \int_{\mathbb{R}^N} (P_t * \mu)(x)g(x)\,dx = \int_{\mathbb{R}^N} g(x)\,d\mu(x)$ for every g in $C_{\text{com}}(\mathbb{R}^N)$.

Problems 6–12 examine the Fourier transform of a measure in $\mathcal{M}^+(\mathbb{R}^N)$, ultimately proving "Bochner's theorem" characterizing the "positive definite functions" on $\mathbb{R}^N$. They take for granted the **Helly–Bray Theorem**, i.e., the statement that if $\{\mu_n\}$ is a sequence in $\mathcal{M}^+(\mathbb{R}^N)$ with $\mu_n(\mathbb{R}^N)$ bounded, then there is a subsequence $\{\mu_{n_k}\}$ convergent to some member μ of $\mathcal{M}^+(\mathbb{R}^N)$ weak-star against $C_{\text{com}}(\mathbb{R}^N)$. The Helly–Bray Theorem will be proved in something like this form in Chapter XI.

6. If μ is in $\mathcal{M}^+(\mathbb{R}^N)$, the **Fourier transform** of μ is defined to be the function $\widehat{\mu}(y) = \int_{\mathbb{R}^N} e^{-2\pi i x\cdot y}\, d\mu(x)$.
   (a) Prove that $\widehat{\mu}$ is bounded and continuous.
   (b) Prove that the Fourier transform of the delta measure at 0 does not vanish at infinity.
   (c) Prove that $\widehat{\mu_1 * \mu_2} = \widehat{\mu_1}\widehat{\mu_2}$ when convolution of finite measures is defined as in Problem 5.
   (d) By forming $\varphi_\varepsilon * \mu$, prove that $\widehat{\mu}$ can equal 0 for some μ in $\mathcal{M}^+(\mathbb{R}^N)$ only if $\mu = 0$.
7. A continuous function $F : \mathbb{R}^N \to \mathbb{C}$ is called **positive definite** if for each finite set of points $x_1, \dots, x_k$ in $\mathbb{R}^N$ and corresponding system of complex numbers $\xi_1, \dots, \xi_k$, the inequality $\sum_{i,j} F(x_i - x_j)\xi_i\overline{\xi_j} \ge 0$ holds. Prove that the continuous function F is positive definite if and only if the inequality $\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} F(x-y)g(x)\overline{g(y)}\,dx\,dy \ge 0$ holds for each member g of $C_{\text{com}}(\mathbb{R}^N)$.
8. Prove that the Fourier transform of any member μ of $\mathcal{M}^+(\mathbb{R}^N)$ is a positive definite function.
9. Using sets of one and then two elements x_i in the definition of positive definite, prove that a positive definite function F must have $F(0) \ge 0$ and $|F(x)| \le F(0)$ for all x.
10. Suppose that F is positive definite, that $\varphi \ge 0$ is in $L^1(\mathbb{R}^N)$, and that $\Phi(x) = \int_{\mathbb{R}^N} e^{2\pi i x\cdot y}\varphi(y)\,dy$. Prove that $F(x)\Phi(x)$ is positive definite.
11. Suppose that F is positive definite. Let $\varepsilon > 0$, and let φ be as in Problem 10, so that $\varphi(x) = \varepsilon^{-N}e^{-\pi\varepsilon^{-2}|x|^2}$ and $\Phi(x) = e^{-\pi\varepsilon^2|x|^2}$.
    (a) The function $F_0(x) = F(x)\Phi(x)$ is positive definite by Problem 10. Prove that it is integrable.
    (b) Using Problem 2 and the alternative definition of positive definite in Problem 7, prove that $\int_{\mathbb{R}^N} \widehat{F_0}(y)|\widehat{g}(y)|^2\,dy \ge 0$ for every function g in $C_{\text{com}}(\mathbb{R}^N)$.
    (c) Deduce from (b) that the function $f_0 = \widehat{F_0}$ is ≥ 0.
    (d) Conclude from (c) that f_0 is integrable with $\int_{\mathbb{R}^N} f_0\,dy = F(0)$, hence that $f_0(y)\,dy$ is a finite Borel measure.
12. **(Bochner's Theorem)** By combining the results of the previous problems with the Helly–Bray Theorem, prove that each positive definite function on $\mathbb{R}^N$ is the Fourier transform of a finite Borel measure.

Problems 13–18 concern a version of the Fourier transform for finite abelian groups, along with the Poisson Summation Formula in that setting. They show for a cyclic group of order $m = pq$ that the use of the idea behind the Poisson Summation Formula makes it possible to compute a Fourier transform in about $pq(p+q)$ steps rather than the expected $m^2 = p^2q^2$ steps. This savings may be iterated in the case of a cyclic group of order 2^n so that the Fourier transform is computed in about $n2^n$

steps rather than the expected 2^{2n} steps. An organized algorithm to implement this method of computation is known as the **fast Fourier transform**.

13. Let G be a finite abelian group. A **multiplicative character** χ of G is a homomorphism of G into the circle group $\{e^{i\theta}\}$. If f and g are two complex-valued functions on G, their L^2 inner product is defined to be $\sum_{t\in G} f(t)\overline{g(t)}$.
    (a) Prove that the set of multiplicative characters of G forms an abelian group under pointwise multiplication, the identity element being the constant function 1 and the inverse of χ being $\overline{\chi}$. This group $\widehat{G}$ is called the **dual group** of G.
    (b) Prove that distinct multiplicative characters are orthogonal and hence that the members of $\widehat{G}$ form a linearly independent set.
    (c) Let J_m be the cyclic group $\{0, 1, 2, \dots, m-1\}$ of integers modulo m under addition, and let $\zeta_m = e^{2\pi i/m}$. For k in J_m define a multiplicative character χ_n of J_m by $\chi_n(k) = (\zeta_m^n)^k$. Prove that the resulting m multiplicative characters exhaust $\widehat{J_m}$ and that $\chi_n\chi_{n'} = \chi_{n+n'}$. Therefore $\widehat{J_m}$ is isomorphic to J_m. For Problems 16–18 below, it will be convenient to identify χ_n with $\chi_n(1) = \zeta_m^n$.
    (d) If G is a direct sum of cyclic groups of orders $m_1, \dots, m_r$, use (c) to exhibit $\prod_{j=1}^r m_j$ distinct members of $\widehat{G}$. Using (b) and the theorem that every finite abelian group is the direct sum of cyclic groups, conclude for any finite abelian group G that these members of $\widehat{G}$ exhaust $\widehat{G}$ and form a basis of $L^2(G)$.

14. Let G be a finite abelian group, and let $\widehat{G}$ be its dual group. The **Fourier transform** of a function f in $L^2(G)$ is the function $\widehat{f}$ on $\widehat{G}$ given by $\widehat{f}(\chi) = \sum_{t\in G} f(t)\overline{\chi(t)}$. Prove that the Fourier transform mapping carries $L^2(G)$ one-one onto $L^2(\widehat{G})$ and that the correct analog of the Fourier inversion formula is $f(t) = |G|^{-1}\sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\chi(t)$, where $|G|$ is the order of G.

15. Let G be a finite abelian group, let H be a subgroup, and let G/H be the quotient group. If t is in G, write $\dot{t}$ for the coset of t in G/H. Let f be in $L^2(G)$ and define $F(\dot{t}) = \sum_{h\in H} f(t+h)$ as a function on G/H. Suppose that χ is a member of $\widehat{G}$ that is identically 1 on H, so that χ descends to a member $\dot{\chi}$ of $\widehat{G/H}$. By imitating steps in the proof of Theorem 8.15, prove that $\widehat{f}(\chi) = \widehat{F}(\dot{\chi})$.

16. Now suppose that $G = J_m$ with $m = pq$; here p and q need not be relatively prime. Let $H = \{0, q, 2q, \dots, (p-1)q\}$ be the subgroup of G isomorphic to J_p, so that $G/H = \{0, 1, 2, \dots, q-1\}$ is isomorphic to J_q. Prove that the characters χ of G identified as in Problem 13c with $\zeta_m^0, \zeta_m^p, \zeta_m^{2p}, \dots, \zeta_m^{(q-1)p}$ are the ones that are identically 1 on H and therefore descend to characters of G/H. Verify that the descended characters $\dot{\chi}$ are the ones identified with $\zeta_q^0, \zeta_q^1, \zeta_q^2, \dots, \zeta_q^{q-1}$. Consequently the formula $\widehat{f}(\chi) = \widehat{F}(\dot{\chi})$ of the previous problem provides a way of computing $\widehat{f}$ at $\zeta_m^0, \zeta_m^p, \zeta_m^{2p}, \dots, \zeta_m^{(q-1)p}$ from the values of $\widehat{F}$. Show that if

$\widehat{F}$ is computed from the definition of Fourier transform in Problem 14, then the number of steps involved in its computation is about q^2, apart from a constant factor. Show therefore that the total number of steps in computing $\widehat{f}$ at these special values of χ is therefore on the order of $q^2 + pq$.

17. In the previous problem show for each k with $0 \le k \le p-1$ that the value of $\widehat{f}$ at $\zeta_m^k, \zeta_m^{p+k}, \zeta_m^{2p+k}, \dots, \zeta_m^{(q-1)p+k}$ can be handled in the same way with a different F by replacing f by a suitable variant of f. Doing so for each k requires p times the number of steps detected in the previous problem, and therefore all of $\widehat{f}$ can be computed in about $p(q^2 + pq) = pq(p+q)$ steps.

18. Show how iteration of this process to compute the Fourier transform of each F, together with further iteration of this process, allows one to compute a Fourier transform for $J_{m_1 m_2 \cdots m_r}$ in about $m_1 m_2 \cdots m_r (m_1 + m_2 + \cdots + m_r)$ steps.

CHAPTER IX

L^p Spaces

Abstract. This chapter extends the theory of the spaces L^1, L^2, and L^∞ to include a whole family of spaces L^p, $1 \leq p \leq \infty$, in order to be able to capture finer quantitative facts about the size of measurable functions and the effect of linear operators on such functions.

Sections 1–2 give the basics about L^p. For general measure spaces these consist of Hölder's inequality, Minkowski's inequality, a completeness theorem, and related results. For Euclidean space they include also facts about convolution.

Sections 3–4 develop some tools that at first may seem quite unrelated to L^p spaces but play a significant role in Section 5. These are the Radon–Nikodym Theorem and two decomposition theorems for additive set functions. The Radon–Nikodym Theorem gives a sufficient condition for writing a measure as a function times another measure.

Section 5 identifies the space of continuous linear functionals on L^p for $1 \leq p < \infty$ when the underlying measure is σ-finite. For one thing this identification makes Alaoglu's Theorem in Chapter V concrete enough so as to be quite useful.

Section 6 discusses the Marcinkiewicz Interpolation Theorem, which allows one to reinterpret suitably bounded operators between two pairs of L^p spaces as bounded between intermediate pairs of L^p spaces as well. The theorem has immediate corollaries for the Hardy–Littlewood maximal function and an approximation to the Hilbert transform, and Section 6 goes on to use each of these corollaries to derive interesting consequences.

1. Inequalities and Completeness

In the context of any measure space, we introduced in Section V.9 the spaces L^1, L^2, and L^∞. Since then, we have used these three spaces to capture quantitative facts about the size of measurable functions. The construction in each case involved introducing a certain pseudonorm in a vector space of functions, thereby making the vector space into a pseudo normed linear space and in particular a pseudometric space. The corresponding metric space obtained from the construction of Proposition 2.12 was L^1, L^2, or L^∞ in the respective cases. For each of the three, the vector-space structure for the pseudometric space yielded a vector-space structure for the metric space, and L^1, L^2, and L^∞ were normed linear spaces. As was true in Chapters V and VI, it continues in the present chapter to be largely a matter of indifference whether the functions in question are real valued or complex valued, hence whether the scalars for these vector spaces are real or complex.

Now we shall enlarge the family consisting of L^1, L^2, L^∞ to a family L^p for $1 \leq p \leq \infty$ in order to be able to capture finer quantitative facts about the size of measurable functions. Enlarging the family in this way makes it possible to get better insight into the behavior of specific operators and to make more helpful estimates with partial differential equations.

Let $(X, \mathcal{A}, \mu)$ be a measure space. We have already dealt with $p = \infty$. For $1 \leq p < \infty$, we consider the set $V = V_p$ of measurable functions f on X such that $\int_X |f|^p \, d\mu$ is finite. This integral is well defined; in fact, f measurable implies $|f|$ measurable, and also, for $c > 0$, $(|f|^p)^{-1}(c, +\infty) = |f|^{-1}(c^{1/p}, +\infty)$. The set V is in fact a vector space of functions. It is certainly closed under scalar multiplication; let us see that it is closed under addition. If f and g are in V, then we have

$$\begin{aligned}(|f(x)| + |g(x)|)^p &\leq \big(\max\{|f(x)|, |g(x)|\} + \max\{|f(x)|, |g(x)|\}\big)^p \\ &= 2^p \max\{|f(x)|^p, |g(x)|^p\} \leq 2^p|f(x)|^p + 2^p|g(x)|^p\end{aligned}$$

for every x in X. Integrating over X, we see that $f + g$ is in V if f and g are in V.

Following the construction of the prototypes L^1 and L^2 in Section V.9, we introduce the expression $\|f\|_p = \big(\int_X |f|^p \, d\mu\big)^{1/p}$ for f in V_p. We would like $\|\cdot\|_p$ to be a pseudonorm in the sense of satisfying

(i) $\|x\|_p \geq 0$ for all $x \in V$,
(ii) $\|cx\|_p = |c|\|x\|_p$ for all scalars c and all $x \in V$,
(iii) $\|x + y\|_p \leq \|x\|_p + \|y\|_p$ for all x and y in V.

Properties (i) and (ii) are certainly satisfied, but a little argument is needed to verify (iii). We return to this matter in a moment. Once the function $\|\cdot\|_p$ on the vector space V_p is known to be a pseudonorm, V_p meets the conditions of being a pseudo normed linear space in the sense of Section V.9.

We can pass to the set of equivalence classes just as in that section, and this set is defined to be L^p or $L^p(X)$ or $L^p(X, \mu)$. The equivalence class of 0 is again the set of all functions vanishing almost everywhere. The function $\|\cdot\|_p$ is well defined on L^p, and L^p is a normed linear space. In particular, it has the structure of a metric space. This handles $1 \leq p < \infty$, and the space L^∞ was constructed in Section V.9.

As is true with L^1, L^2, and L^∞, one sometimes relaxes the terminology and works with the members of $L^p(X)$ as if they were functions, saying, "Let the function f be in $L^p(X)$" or "Let f be an L^p function." There is little possibility of ambiguity in using such expressions.

Let us return to property (iii) above. This will be proved as Minkowski's inequality below. But first we prove a numerical lemma and then "Hölder's inequality," which is a version for L^p of the Schwarz inequality for L^2. Hölder's inequality makes use of the **dual index** p' to p, defined by the equality $\frac{1}{p} + \frac{1}{p'} = 1$. The dual index to 1 is ∞, and vice versa. The index 2 is its own dual.

Lemma 9.1. If s, t, α, and β are real numbers ≥ 0 with $\alpha + \beta = 1$, then

$$s^\alpha t^\beta \leq \alpha s + \beta t.$$

PROOF. If any of s, t, α, β is 0, the result is certainly true. If all are nonzero, consider the function

$$f(x) = \alpha x^{\alpha-1} + (1-\alpha)x^\alpha,$$

defined for $x > 0$. The derivative $f'(x) = (1-\alpha)\alpha x^{\alpha-2}(x-1)$ is < 0 for $0 < x < 1$, is $= 0$ for $x = 1$, and is > 0 for $x > 1$. Therefore $f(x)$ assumes its absolute minimum value for $x = 1$. Since $f(1) = 1$, we have

$$1 \leq \alpha x^{\alpha-1} + (1-\alpha)x^\alpha = \alpha x^{-\beta} + \beta x^\alpha$$

for all $x > 0$. The lemma follows by putting $x = t/s$ in this inequality and by multiplying both sides by $s^\alpha t^\beta$.

REMARK. Alternatively, this lemma can be proved by Lagrange multipliers in the same way that Problem 20 at the end of Chapter III suggested using for the arithmetic-geometric mean inequality.

Theorem 9.2 (Hölder's inequality). Let $(X, \mathcal{A}, \mu)$ be any measure space, let $1 \leq p \leq \infty$, and let p' be the dual index to p. If f is in L^p and g is in $L^{p'}$, then fg is in L^1, and

$$\|fg\|_1 \leq \|f\|_p \|g\|_{p'}.$$

REMARK. The inequality holds trivially if $\|f\|_p = +\infty$ or $\|g\|_{p'} = +\infty$.

PROOF. We already know the result if $p = 1$ and $p' = \infty$ or the other way around. Thus suppose that $p > 1$ and $p' > 1$. We may assume that neither f nor g is 0 almost everywhere. Then we can apply Lemma 9.1 with $\alpha = p^{-1}$, $\beta = p'^{-1}$,

$$s = \frac{|f(x)|^p}{\int_X |f|^p \, d\mu}, \qquad \text{and} \qquad t = \frac{|g(x)|^{p'}}{\int_X |g|^{p'} \, d\mu},$$

getting

$$\frac{|f(x)g(x)|}{\|f\|_p \|g\|_{p'}} \leq \frac{|f(x)|^p}{p \int_X |f|^p \, d\mu} + \frac{|g(x)|^{p'}}{p' \int_X |g|^{p'} \, d\mu}.$$

Integrating, we obtain

$$\frac{\int_X |fg| \, d\mu}{\|f\|_p \|g\|_{p'}} \leq \frac{1}{p} + \frac{1}{p'} = 1,$$

and the conclusions of the theorem follow.

Theorem 9.3 (Minkowski's inequality). Let $(X, \mathcal{A}, \mu)$ be any measure space, and let $1 \le p \le \infty$. If f and g are in L^p, then $f+g$ is in L^p and

$$\|f+g\|_p \le \|f\|_p + \|g\|_p.$$

REMARK. The theorem assumes the usual convention that $f+g$ is made to be 0 at any point x where $f(x)+g(x)$ is not defined. The set where this change occurs is of measure 0 since f and g have to be finite almost everywhere to be in L^p.

PROOF. We have already seen that $f+g$ is in L^p, and we know the inequality for $p=1$ and $p=\infty$ from Section V.9. For $1<p<\infty$, let p' be the dual index. We apply Hölder's inequality (Theorem 9.2) to f and $|f+g|^{p-1}$ and to g and $|f+g|^{p-1}$ to obtain

$$\begin{aligned}
\textstyle\int_X |f+g|^p\,d\mu &\le \textstyle\int_X |f+g|\,|f+g|^{p-1}\,d\mu \\
&\le \textstyle\int_X |f|\,|f+g|^{p-1}\,d\mu + \int_X |g|\,|f+g|^{p-1}\,d\mu \\
&\le \|f\|_p\big(\textstyle\int |f+g|^{(p-1)p'}\,d\mu\big)^{1/p'} + \|g\|_p\big(\textstyle\int |f+g|^{(p-1)p'}\,d\mu\big)^{1/p'} \\
&= \big(\textstyle\int_X |f+g|^p\,d\mu\big)^{1/p'}(\|f\|_p + \|g\|_p),
\end{aligned}$$

the last step holding because $(p-1)p'=p$. If $\|f+g\|_p=0$, the inequality of the theorem is certainly true. Otherwise the inequality of the theorem follows after dividing the inequality of the display by $\big(\int_X |f+g|^p\,d\mu\big)^{1/p'}$, which we know to be finite, and using the fact that $1-\frac{1}{p'}=\frac{1}{p}$.

Thus L^p is a normed linear space for $1 \le p \le \infty$. Let us derive some of its properties.

Proposition 9.4. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $1 \le p < \infty$. Then every indicator function of a set of finite measure is in $L^p(X)$, and the smallest closed subspace of $L^p(X)$ containing all such indicator functions is $L^p(X)$ itself. Consequently

(a) the set of simple functions built from sets of finite measure lies in every $L^p(X)$ for $1 \le p \le \infty$ and is dense in $L^p(X)$ if $1 \le p < \infty$,

(b) $1 \le p_1 \le p \le p_2 \le \infty$ and $p<\infty$ together imply that $L^{p_1}(X) \cap L^{p_2}(X)$ is dense in $L^p(X)$.

In addition,

(c) $1 \le p_1 \le p \le p_2 \le \infty$ implies that $L^p(X) \subseteq L^{p_1}(X) + L^{p_2}(X)$.

PROOF. The conclusion in the second sentence of the proposition is proved by the same argument as for Proposition 5.56. Part (a) then follows from Proposition 5.55d. Part (b) follows by combining these two results once it is known that $L^{p_1}(X) \cap L^{p_2}(X) \subseteq L^p(X)$. For this inclusion let f be in $L^{p_1}(X) \cap L^{p_2}(X)$. We may assume that $p < \infty$. If $p_2 < \infty$, then

$$\begin{aligned}\textstyle\int_X |f|^p\,d\mu &= \textstyle\int_{\{|f|>1\}} |f|^p\,d\mu + \int_{\{|f|\le 1\}} |f|^p\,d\mu \\ &\le \textstyle\int_{\{|f|>1\}} |f|^{p_2}\,d\mu + \int_{\{|f|\le 1\}} |f|^{p_1}\,d\mu < +\infty,\end{aligned}$$

and hence f is in $L^p(X)$. If $p_2 = \infty$, then $\{|f| > 1\}$ has finite measure since f is in L^{p_1} and $p_1 < \infty$. Thus

$$\begin{aligned}\textstyle\int_X |f|^p\,d\mu &= \textstyle\int_{\{|f|>1\}} |f|^p\,d\mu + \int_{\{|f|\le 1\}} |f|^p\,d\mu \\ &\le \|f\|_\infty^p\,\mu(\{|f| > 1\}) + \textstyle\int_{\{|f|\le 1\}} |f|^{p_1}\,d\mu < +\infty,\end{aligned}$$

and again f is in $L^p(X)$. This completes the proof of (b).

For (c), let f be in L^p, and write $f = f_1 + f_2$, where

$$f_1(x) = \left\{\begin{matrix} f(x) & \text{if } |f(x)| > 1 \\ 0 & \text{otherwise}\end{matrix}\right\} \quad \text{and} \quad f_2(x) = \left\{\begin{matrix} f(x) & \text{if } |f(x)| \le 1 \\ 0 & \text{otherwise}\end{matrix}\right\}.$$

Then

$$\int_X |f_1|^{p_1}\,d\mu = \int_{\{|f|>1\}} |f|^{p_1}\,d\mu \le \int_{\{|f|>1\}} |f|^p\,d\mu < \infty$$

shows that f_1 is in $L^{p_1}(X)$. It is apparent that f_2 is in $L^\infty(X)$, and thus f_2 is certainly in $L^{p_2}(X)$ if $p_2 = \infty$. If $p_2 < \infty$, then

$$\int_X |f_2|^{p_2}\,d\mu = \int_{\{|f|\le 1\}} |f|^{p_2}\,d\mu \le \int_{\{|f|\le 1\}} |f|^p\,d\mu < \infty$$

shows that f_2 is in $L^{p_2}(X)$. This proves (c).

Hölder's inequality allows us to prove the following supplement to the conclusions of Proposition 9.4.

Proposition 9.5. Let $(X, \mathcal{A}, \mu)$ be any measure space. Let $1 \le p_1 < p < p_2$, and define t with $0 \le t \le 1$ by $\frac{1}{p} = \frac{1-t}{p_1} + \frac{t}{p_2}$. Then

$$\|f\|_p \le \|f\|_{p_1}^{1-t}\|f\|_{p_2}^t.$$

PROOF. First suppose that $p_2 < \infty$. Since $\frac{1}{p} > \frac{1-t}{p_1}$, we can find b with $1 < b < +\infty$ such that $\frac{1}{bp} = \frac{1-t}{p_1}$. If b' denotes the dual index, then $\frac{1}{b'p} =$

$\frac{1}{p} - \frac{1}{bp} = \frac{1}{p} - \frac{1-t}{p_1} = \frac{t}{p_2}$. Define a by the equation $ab = p_1$. Then $(p-a)b' = \big(p - \frac{p_1}{b}\big)\frac{p_2}{tp} = p_2\big(\frac{1}{t} - \frac{p_1}{btp}\big) = p_2\big(\frac{1}{t} - \frac{1-t}{t}\big) = p_2$.

We write $|f|^p = |f|^a|f|^{p-a}$. Application of Hölder's inequality with index b and dual index b' gives $\int |f|^p\,d\mu \le \big(\int |f|^{ab}\,d\mu\big)^{1/b}\big(\int |f|^{(p-a)b'}\,d\mu\big)^{1/b'}$, and hence

$$\|f\|_p \le \big(\textstyle\int |f|^{ab}\,d\mu\big)^{1/(bp)}\big(\textstyle\int |f|^{(p-a)b'}\,d\mu\big)^{1/(b'p)}.$$

We have seen that $ab = p_1$, $1/(bp) = (1-t)/p_1$, $(p-a)b' = p_2$, and $1/(b'p) = t/p_2$. Thus the inequality reads $\|f\|_p \le \|f\|_{p_1}^{1-t}\|f\|_{p_2}^t$, and the proof is complete when $p_2 < \infty$.

When $p_2 = \infty$, we write $|f|^p = |f|^{p_1}|f|^{p-p_1}$. Replacing $|f|^{p-p_1}$ by its essential supremum gives $\int |f|^p\,d\mu \le \|f\|_\infty^{p-p_1}\int |f|^{p_1}\,d\mu$ and hence $\|f\|_p$ is

$$\le \big(\textstyle\int |f|^{p_1}\,d\mu\big)^{1/p}\|f\|_\infty^{(p-p_1)/p} = \big(\textstyle\int |f|^{p_1}\,d\mu\big)^{(1-t)/p_1}\|f\|_\infty^{1-p_1/p} = \|f\|_{p_1}^{1-t}\|f\|_\infty^t.$$

This completes the proof when $p_2 = \infty$.

We have already made serious use of the completeness of L^p for p equal to 1, 2, and ∞ as proved in Theorem 5.58. As might be expected, this result extends to be valid for the other values of p.

Theorem 9.6. Let $(X, \mathcal{A}, \mu)$ be any measure space, and let $1 \le p \le \infty$. Any Cauchy sequence $\{f_k\}$ in L^p has a subsequence $\{f_{k_n}\}$ such that $\|f_{k_n} - f_{k_m}\|_p \le C_{\min\{m,n\}}$ with $\sum_n C_n < +\infty$. A subsequence $\{f_{k_n}\}$ with this property is necessarily Cauchy pointwise almost everywhere. If f denotes the almost-everywhere limit of $\{f_{n_k}\}$, then the original sequence $\{f_k\}$ converges to f in L^p. Consequently the space L^p, when regarded as a metric space, is complete in the sense that every Cauchy sequence converges.

REMARK. As in the case with p equal to $1, 2$, and ∞, the detail is important. The detailed statement of the theorem allows us to conclude, among other things, that if a sequence of functions is convergent in L^{p_1} and in L^{p_2}, then the limit functions in the two spaces are equal almost everywhere.

PROOF. We may assume that $p < \infty$, the case $p = \infty$ having been handled in Theorem 5.58. The argument for $1 \le p < \infty$ is word-for-word the same as in the proof for $p = 1$ and $p = 2$ of Theorem 5.58.

In Section V.9 the inequality $\|f + g\|_p \le \|f\|_p + \|g\|_p$ for p equal to $1, 2$, or ∞ says in words that "the norm of a sum is $\le$ the sum of the norms." In that section we obtained a generalization for those values of p, saying that "the norm of an integral is $\le$ the integral of the norms." The generalization continues to be valid for the other p's under study; the proof amounts to a direct derivation from Hölder's inequality.

Theorem 9.7 (Minkowski's inequality for integrals). Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let $1 \leq p \leq \infty$. If f is measurable on $X \times Y$, then

$$\Big\| \int_X f(x, y)\, d\mu(x) \Big\|_{p, d\nu(y)} \leq \int_X \| f(x, y) \|_{p, d\nu(y)}\, d\mu(x)$$

in the following sense: The integrand on the right side is measurable. If the integral on the right is finite, then for almost every y $[d\nu]$ the integral on the left is defined; when it is redefined to be 0 for the exceptional y's, then the formula holds.

PROOF. Theorem 5.60 handles $p = 1$ and $p = \infty$, and we may assume that $1 < p < \infty$. The measurability question is handled for $1 < p < \infty$ in the same way as in Theorem 5.60 for $p = 2$. In proving the inequality, we may assume without loss of generality that $f \geq 0$. The generalization of the computation in the proof of Theorem 9.3 makes use of Fubini's Theorem and proceeds as follows:

$$\begin{aligned}
&\textstyle\int_Y \big| \int_X f(x, y)\, d\mu(x) \big|^p\, d\nu(y) \\
&\quad = \textstyle\int_Y \big| \int_X f(x, y)\, d\mu(x) \big| \big| \int_X f(x', y)\, d\mu(x') \big|^{p-1}\, d\nu(y) \\
&\quad = \textstyle\int_X \big\{ \int_Y f(x, y) \, \big| \int_X f(x', y)\, d\mu(x') \big|^{p-1}\, d\nu(y) \big\}\, d\mu(x) \\
&\quad \leq \textstyle\int_X \big(\int_Y |f(x, y)|^p\, d\nu(y) \big)^{1/p} \\
&\qquad\quad \times \big(\textstyle\int_Y \big| \int_X f(x', y)\, d\mu(x') \, \big|^{(p-1)p'}\, d\nu(y) \big)^{1/p'}\, d\mu(x) \\
&\quad = \big(\textstyle\int_X \| f(x, y) \|_{p, d\nu(y)}\, d\mu(x) \big) \big(\int_Y \big| \int_X f(x', y)\, d\mu(x') \big|^p\, d\nu(y) \big)^{1/p'}.
\end{aligned}$$

The next-to-last step uses Hölder's inequality (Theorem 9.2), and the last step uses the fact that $(p - 1)p' = p$.

In order to complete the proof, we need to be able to divide by the factor $\big(\int_Y \big| \int_X f(x', y)\, d\mu(x') \big|^p\, d\nu(y) \big)^{1/p'}$. There is no problem with the theorem if this factor is 0, since then the left side of the inequality of the theorem is 0. A problem occurs if this factor is infinite. Instead of trying to prove directly that this factor is finite (and hence the division is allowable), let us retreat to the special case that f is bounded and is equal to 0 off an abstract rectangle of finite $\mu \times \nu$ measure. Then the factor in question is certainly finite, the division is allowable, and we obtain the inequality of the theorem. To handle general measurable $f \geq 0$, we do not attempt to justify this division. Instead, we observe that the validity of the inequality in the theorem when f is bounded and is equal to 0 off a set of finite $\mu \times \nu$ measure implies the validity of the inequality in general, by a routine application of monotone convergence. This completes the proof.

The last basic fact about L^p spaces is the identification of continuous linear functionals on L^p, at least when p is finite. Deriving the necessary tools for this analysis will require a digression, and we shall return to this topic in Section 5. Meanwhile, we can easily obtain one part of the identification of continuous linear functionals, as in Proposition 9.8 below. It amounts to a combination of Hölder's inequality and a converse, and it gives a way of computing L^p norms by starting with computations that are linear.

Proposition 9.8. Let $(X, \mathcal{A}, \mu)$ be any measure space, let $1 \leq p \leq \infty$, and let p' be the dual index. If $p < \infty$, then

$$\|f\|_p = \sup_{\substack{g \in L^{p'}, \\ \|g\|_{p'} \leq 1}} \left| \int_X fg \, d\mu \right|,$$

and this equality remains valid for $p = \infty$ if μ is σ-finite.

REMARK. The equality can fail when $p = \infty$ and μ is not σ-finite. Problem 4 at the end of the chapter gives an example.

PROOF. With $1 \leq p \leq \infty$, if g is in $L^{p'}$ with $\|g\|_{p'} \leq 1$, then Hölder's inequality gives $\left|\int fg \, d\mu\right| \leq \int |fg| \, d\mu \leq \|f\|_p \|g\|_{p'} \leq \|f\|_p$. Taking the supremum over g with $\|g\|_{p'} \leq 1$ shows that $\sup_g \left|\int fg \, d\mu\right| \leq \|f\|_p$.

For the reverse inequality we may assume that $\|f\|_p \neq 0$. First suppose that $1 < p < \infty$. Define $g(x)$ by

$$g(x) = \begin{cases} \|f\|_p^{-(p-1)} \, \overline{f(x)} \, |f(x)|^{p-2} & \text{if } f(x) \neq 0, \\ 0 & \text{if } f(x) = 0. \end{cases}$$

Then $\int |g(x)|^{p'} \, d\mu = \|f\|_p^{-(p-1)p'} \int |f(x)|^{(p-1)p'} \, d\mu = \|f\|_p^{-p} \int |f(x)|^p \, d\mu = 1$. For this g, we have $\left|\int fg \, d\mu\right| = \|f\|_p^{-(p-1)} \int |f|^p \, d\mu = \|f\|_p$. Thus the supremum over the relevant g's of $\left|\int fg \, d\mu\right|$ is $\geq \|f\|_p$.

Next suppose that $p = 1$. If we define $g(x)$ to be $\overline{f(x)}/|f(x)|$ when $f(x) \neq 0$ and to be 0 when $f(x) = 0$, then $\|g\|_\infty = 1$ and $\left|\int fg \, d\mu\right| = \int |f|^2/|f| \, d\mu = \|f\|_1$, and the supremum over g of $\left|\int fg \, d\mu\right|$ is $\geq \|f\|_1$.

Finally suppose that $p = \infty$. Let $\epsilon > 0$ be given with $\epsilon \leq \|f\|_\infty$, and let E be the set where $|f(x)| \geq \|f\|_\infty - \epsilon$. Since μ is σ-finite, there must exist a subset of E with nonzero finite measure. If F is such a subset and if $g(x)$ is defined to be $\mu(F)^{-1}\overline{f(x)}/|f(x)|$ when x is in F and to be 0 when x is in F^c, then $\|g\|_1 = 1$ and $\left|\int_X fg \, d\mu\right| = \mu(F)^{-1} \int_F |f| \, d\mu \geq \|f\|_\infty - \epsilon$. Thus the supremum over g of $\left|\int_X fg \, d\mu\right|$ is $\geq \|f\|_\infty - \epsilon$. Since ϵ is arbitrary, the supremum over g of $\left|\int_X fg \, d\mu\right|$ is $\geq \|f\|_\infty$.

2. Convolution Involving L^p

In this section we collect results about L^p spaces that extend facts proved about L^1, L^2, and L^∞ in the first three sections of Chapter VI.

Proposition 9.9. If μ is a Borel measure on a nonempty open set V in $\mathbb{R}^N$ and if $1 \le p < \infty$, then

(a) $C_{\text{com}}(V)$ is dense in $L^p(V, \mu)$,

(b) the smallest closed subspace of $L^p(V, \mu)$ containing all indicator functions of compact subsets of V is $L^p(V, \mu)$ itself,

(c) $L^p(V, \mu)$ is separable.

PROOF. Parts (a) and (b) are proved from Lemma 6.22c, the regularity of μ (Theorem 6.25), Proposition 9.4, and Proposition 5.56 by the same kind of argument as for Corollary 6.4. Part (c) is obtained as a consequence in the same way that Corollary 6.27d follows from the other parts of that corollary.

The remaining results in this section concern Lebesgue measure in $\mathbb{R}^N$, and the L^p spaces are understood to be $L^p(\mathbb{R}^N, \{\text{Borel sets}\}, dx)$.

Proposition 9.10. Let $1 < p < \infty$, and let p' be the dual index. Convolution is defined in the following additional cases beyond those listed in Proposition 6.14, and the indicated inequalities hold:

(e) for f in $L^1(\mathbb{R}^N, dx)$ and g in $L^p(\mathbb{R}^N, dx)$, and then $\|f*g\|_p \le \|f\|_1\|g\|_p$,
for f in $L^p(\mathbb{R}^N, dx)$ and g in $L^1(\mathbb{R}^N, dx)$, and then $\|f*g\|_p \le \|f\|_p\|g\|_1$,

(f) for f in $L^p(\mathbb{R}^N, dx)$ and g in $L^{p'}(\mathbb{R}^N, dx)$, and then $\|f * g\|_\infty \le \|f\|_p\|g\|_{p'}$,
for f in $L^{p'}(\mathbb{R}^N, dx)$ and g in $L^p(\mathbb{R}^N, dx)$, and then $\|f * g\|_\infty \le \|f\|_{p'}\|g\|_p$.

PROOF. The two conclusions in (e) follow from Minkowski's inequality for integrals (Theorem 9.7) in the same way that the special case of $p = 2$ was proved in Proposition 6.14 from Theorem 5.60. The two conclusions in (f) follow from Hölder's inequality (Theorem 9.2) in the same way that the special case $p = p' = 2$ was proved in Proposition 6.14 from the Schwarz inequality.

Proposition 9.11. If $1 \le p < \infty$, then translation of a function is continuous in the translation parameter in $L^p(\mathbb{R}^N, dx)$. In other words, if f is in $L^p(\mathbb{R}^N, dx)$, then $\lim_{h\to 0} \|\tau_{t+h} f - \tau_t\|_p = 0$ for all t.

PROOF. This follows from the denseness of $C_{\text{com}}(\mathbb{R}^N)$ in $L^p(\mathbb{R}^N, dx)$ (Proposition 9.9a) and is proved in the same way that Proposition 6.16 is derived from Corollary 6.4a.

Proposition 9.12. Let $1 \le p \le \infty$, and let p' be the dual index. Then the convolution of an L^p function with an $L^{p'}$ function results in an everywhere-defined bounded uniformly continuous function, not just an L^∞ function. Moreover,

$$\|f * g\|_{\sup} \le \|f\|_p \|g\|_{p'}.$$

PROOF. This extends Proposition 6.18 and is derived for $1 < p < \infty$ from Propositions 9.10 and 9.11 in the same way that Proposition 6.18 is derived for $p = 2$ from Propositions 6.14 and 6.16.

Theorem 9.13. Let φ be in $L^1(\mathbb{R}^N, dx)$, define

$$\varphi_\varepsilon(x) = \varepsilon^{-N}\varphi(\varepsilon^{-1}x) \qquad \text{for } \varepsilon > 0,$$

and put $c = \int_{\mathbb{R}^N} \varphi(x)\,dx$. If f is in $L^p(\mathbb{R}^N, dx)$ with $1 \le p < \infty$, then

$$\lim_{\varepsilon \downarrow 0} \|\varphi_\varepsilon * f - cf\|_p = 0.$$

PROOF. This is derived from Minkowski's inequality for integrals (Theorem 9.7) and the continuity of translation in L^p (Proposition 9.11) in the same way that Theorem 6.20a is derived for $p = 2$ from Theorem 5.60 and Proposition 6.16.

3. Jordan and Hahn Decompositions

Now we digress before returning in Section 5 to the subject of continuous linear functionals on L^p spaces. The subject of the present section is decompositions of additive and completely additive real-valued set functions into positive and negative parts. This material will be applied in Section 4 to obtain the Radon–Nikodym Theorem, an abstract generalization of some consequences of Lebesgue's theory of differentiation of integrals. In turn, we shall use the Radon–Nikodym Theorem in Section 5 to address the subject of continuous linear functionals on L^p spaces.

A real-valued additive set function ν on an algebra of sets is said to be **bounded** if $|\nu(E)| \le C$ for all E in the algebra. A real-valued completely additive set function on a σ-algebra of sets is said to be a **signed measure**.

Theorem 9.14 (Jordan decomposition). Let ν be a bounded additive set function on an algebra $\mathcal{A}$ of sets, and define set functions ν^+ and ν^- on $\mathcal{A}$ by

$$\nu^+(E) = \sup_{\substack{F \subseteq E,\\ F \in \mathcal{A}}} \nu(F) \qquad \text{and} \qquad \nu^-(E) = -\inf_{\substack{F \subseteq E,\\ F \in \mathcal{A}}} \nu(F).$$

Then ν^+ and ν^- are nonnegative bounded additive set functions on $\mathcal{A}$ such that $\nu = \nu^+ - \nu^-$. They are completely additive if ν is completely additive. In any event, the decomposition $\nu = \nu^+ - \nu^-$ is minimal in the sense that an equality $\nu = \mu^+ - \mu^-$ in which μ^+ and μ^- are nonnegative bounded additive set functions must have $\nu^+ \le \mu^+$ and $\nu^- \le \mu^-$.

PROOF. First let us see that ν^+ is additive always. In fact, let E_1 and E_2 be disjoint members of $\mathcal{A}$. If $F \subseteq E_1 \cup E_2$, then the additivity of ν implies that $\nu(F) = \nu(F \cap E_1) + \nu(F \cap E_2) \le \nu^+(E_1) + \nu^+(E_2)$. Hence

$$\nu^+(E_1 \cup E_2) \le \nu^+(E_1) + \nu^+(E_2).$$

On the other hand, if $F_1 \subseteq E_1$ and $F_2 \subseteq E_2$, then $\nu(F_1) + \nu(F_2) = \nu(F_1 \cup F_2) \le \nu^+(E_1 \cup E_2)$. Taking the supremum over F_1 and then over F_2 gives

$$\nu^+(E_1) + \nu^+(E_2) \le \nu^+(E_1 \cup E_2).$$

Thus ν^+ is additive.

Second let us see that ν^+ is completely additive if ν is completely additive. Let E_n be a disjoint sequence of sets in $\mathcal{A}$ whose union E is in $\mathcal{A}$. If $F \subseteq E$, then the complete additivity of ν implies that $\nu(F) = \sum_n \nu(F \cap E_n) \le \sum_n \nu^+(E_n)$. Hence $\nu^+(E) \le \sum_{n=1}^\infty \nu^+(E_n)$. On the other hand, the fact that ν^+ is nonnegative additive implies for every N that $\sum_{n=1}^N \nu^+(E_n) = \nu^+(E_1 \cup \cdots \cup E_N) \le \nu^+(E)$. Thus $\sum_{n=1}^\infty \nu^+(E_n) \le \nu^+(E)$. Therefore ν^+ is completely additive.

Third let us see that $\nu = \nu^+ - \nu^-$. This equality will imply also that ν^- is additive and that ν^- is completely additive if ν is completely additive. Form $\nu(E) + \nu^-(E) = \nu(E) + \sup_{F \subseteq E}\{-\nu(F)\}$; we are to show that this equals $\nu^+(E)$. For any $F \subseteq E$, we have $\nu(E) + (-\nu(F)) = \nu(E - F) \le \nu^+(E)$. Taking the supremum over F gives $\nu(E) + \nu^-(E) \le \nu^+(E)$. In the reverse direction, $F \subseteq E$ implies that $\nu(F) = \nu(E) - \nu(E - F) \le \nu(E) + \sup_{G \subseteq E}\{-\nu(G)\} = \nu(E) + \nu^-(E)$. Taking the supremum over F gives $\nu^+(E) \le \nu(E) + \nu^-(E)$. This proves the decomposition $\nu = \nu^+ - \nu^-$.

Finally we prove the minimality of the decomposition. Let $\nu = \mu^+ - \mu^-$ with μ^+ and μ^- nonnegative additive. If $F \subseteq E$, then we can write $\nu(F) = \mu^+(F) - \mu^-(F) \le \mu^+(F) \le \mu^+(E)$. Taking the supremum over F gives $\nu^+(E) \le \mu^+(E)$. Similarly $\nu^- \le \mu^-$.

Theorem 9.15 (Hahn decomposition). If ν is a bounded signed measure on a σ-algebra $\mathcal{A}$ of subsets of X, then there exist disjoint measurable sets P and N in $\mathcal{A}$ with $X = P \cup N$ such that $\nu(E) \ge 0$ for all sets $E \subseteq P$ and $\nu(E) \le 0$ for all sets $E \subseteq N$.

PROOF. Write $\nu = \nu^+ - \nu^-$ as in Theorem 9.14. If $\epsilon > 0$ is given, choose A in $\mathcal{A}$ with $\nu(A) \geq \nu^+(X) - \epsilon$. Then

$$\nu^-(A) = \nu^+(A) - \nu(A) \leq \nu^+(A) - \nu^+(X) + \epsilon \leq \epsilon$$

and
$$\nu^+(A^c) = \nu^+(X) - \nu^+(A) \leq \nu(A) + \epsilon - \nu^+(A) \leq \epsilon.$$

By taking $P_0 = A$ and $N_0 = A^c$, we see that for any $\epsilon > 0$ we can write $X = P_0 \cup N_0$ disjointly with $\nu^+(N_0) \leq \epsilon$ and $\nu^-(P_0) \leq \epsilon$.

For $n \geq 1$, write $X = P_n \cup N_n$ disjointly with $\nu^+(N_n) \leq 2^{-n}$ and $\nu^-(P_n) \leq 2^{-n}$. Define

$$P = \bigcup_{n=1}^\infty \bigcap_{m=n}^\infty P_m \qquad \text{and} \qquad N = P^c = \bigcap_{n=1}^\infty \bigcup_{m=n}^\infty N_m.$$

These sets are in $\mathcal{A}$ since $\mathcal{A}$ is a σ-algebra. Theorem 9.14 shows that ν^- is completely additive, and hence $\nu^-(P) \leq \sum_{n=1}^\infty \nu^-\big(\bigcap_{m=n}^\infty P_m\big)$. The right side is 0 since $\nu^-\big(\bigcap_{m=n}^\infty P_m\big) \leq \nu^-(P_{n+k}) \leq 2^{-(n+k)}$ for all $k \geq 0$, and therefore $\nu^-(P) = 0$. In addition, every n has $\nu^+(N) \leq \nu^+\big(\bigcup_{m=n}^\infty \nu^+(N_m)\big) \leq \sum_{m=n}^\infty \nu^+(N_m) \leq \sum_{m=n}^\infty 2^{-m} = 2^{-n+1}$, and therefore $\nu^+(N) = 0$.

4. Radon–Nikodym Theorem

The Lebesgue decomposition of Chapter VII says that any Stieltjes measure μ on the line decomposes as $\mu(E) = \int_E f\,dx + \mu_s$ with μ_s concentrated on a Borel set of Lebesgue measure 0. The function f is obtained in that chapter as the derivative almost everywhere of the distribution function of μ, hence as the limit of $\mu(I)/m(I)$ as intervals I shrink to a point; here m is Lebesgue measure. In this formulation of the result, the geometry of the line plays an essential role, and attempts to generalize to abstract settings the construction of f from limits of $\mu(I)/m(I)$ have not been fruitful.

Nevertheless, the Lebesgue decomposition itself turns out to be a general measure-theory theorem, valid for any two measures in place of μ and dx, as long as suitable finiteness conditions are satisfied. For a reinterpretation of the results of Chapter VII, the heart of the matter is that one can tell in advance which μ's have $\mu(E) = \int_E f\,dx$ with the singular term μ_s absent. The answer is given by the equivalent conditions of Proposition 7.11, which are taken in that chapter as a definition of "absolute continuity" of μ with respect to dx. The remarkable fact is that those conditions continue to be equivalent when any two finite measures replace μ and dx. This is the content of the Radon–Nikodym Theorem, which we shall prove in this section, and then a version of the Lebesgue decomposition will follow as a consequence.

Let X be a nonempty set, and let $\mathcal{A}$ be a σ-algebra of subsets of X. If μ and ν are measures defined on $\mathcal{A}$, we say that ν is **absolutely continuous** with respect to μ, written $\nu \ll \mu$, if $\nu(E) = 0$ whenever $\mu(E) = 0$.

Theorem 9.16 (Radon–Nikodym Theorem). Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space, and let ν be a σ-finite measure on $\mathcal{A}$ with $\nu \ll \mu$. Then there exists a measurable $f \geq 0$ such that $\nu(E) = \int_E f\, d\mu$ for all E in $\mathcal{A}$, and f is unique up to a set of μ measure 0.

The Radon–Nikodym Theorem has two chief initial applications. One is to the identification of continuous linear functionals on L^p for $1 \leq p < \infty$, and the other is to the construction of "conditional expectation" in probability theory. The application to L^p will be given in Section 5, and the application to conditional expectation appears in Problems 23–26 at the end of the chapter.

In both applications one needs a version of the theorem in which the completely additive set function ν is complex-valued but not necessarily ≥ 0. We take up this extension of the theorem later in this section.

Most of the effort in the proof goes into showing existence when μ and ν are both finite measures, as we shall see. In this setting we can quickly use the Hahn decomposition (Theorem 9.15) to get an idea how to construct f: Imagine that $\nu(E) = \int_E f\, d\mu$ for all E. Fix c and d, and let S be the set of x's where $c \leq f(x) < d$. On any subset E of S, we then have $c\mu(E) \leq \nu(E) \leq d\mu(E)$. In other words, the bounded signed measure $\nu - c\mu$ is ≥ 0 on every subset of S, and the bounded signed measure $\nu - d\mu$ is ≤ 0 on every subset of S. Let $X = P_c \cup N_c$ and $X = P_d \cup N_d$ be Hahn decompositions of $\nu - c\mu$ and $\nu - d\mu$ with respect to μ. Then it is reasonable to expect S to be $P_c \cap N_d$. In particular, c is a good lower bound for the values of f on S. It is easy to imagine that we can use this process repeatedly to obtain a monotone sequence of functions $f_n \geq 0$ tending to the desired function f.

Actually, this argument can be pushed through, but handling the details is a good deal more complicated than one might at first suppose. The reason is that a Hahn decomposition is not necessarily unique. Sets of measure 0 account for the nonuniqueness, and the particular measures yielding these sets of measure 0 are constantly changing. The complication is that one has to adjust all the Hahn decompositions to satisfy various compatibility conditions. We shall not pursue this idea because a simpler proof is available.

Proof of uniqueness in Theorem 9.16. Suppose that f and g are nonnegative measurable functions with $\int_E f\, d\mu = \int_E g\, d\mu$ for every measurable E. If F is a set where the equal integrals $\int_F f\, d\mu$ and $\int_F g\, d\mu$ are finite, then $\int_{E\cap F} (f - g)\, d\mu = 0$ for every measurable subset $E \cap F$ of F. If E is taken as the set where $f > g$, then Corollary 5.23 shows that $f = g$ a.e. on $E \cap F$. Similarly $f = g$ a.e. on the set $E^c \cap F$, where $f \leq g$. Thus $f = g$ a.e. on F. By σ-finiteness of μ and ν, we can write $X = \bigcup_{n=1}^{\infty} X_n$ disjointly with $\mu(X_n)$ and $\nu(X_n)$ finite for all n. Taking F equal to each X_n in turn, we see that $f = g$ a.e. on each X_n, and we conclude that $f = g$ a.e. on X.

PROOF OF EXISTENCE IN THEOREM 9.16 WHEN μ AND ν ARE FINITE. Let $\mathcal{F}(\nu)$ be the set of all $f \geq 0$ in $L^1(X, \mu)$ such that $\int_E f\,d\mu \leq \nu(E)$ for all sets E in $\mathcal{A}$. The zero function is in $\mathcal{F}(\nu)$, and thus it makes sense to define

$$C = \sup_{f \in \mathcal{F}(\nu)} \int_X f\,d\mu.$$

Let $\{f_n\}$ be a sequence in $\mathcal{F}(\nu)$ with $\lim_n \int_X f_n\,d\mu = C$.

Let us show that there is no loss of generality in assuming that the f_n satisfy $f_1 \leq f_2 \leq \cdots$. To show this, it is enough to show that g and h in $\mathcal{F}(\nu)$ implies that $\max\{g, h\}$ is in $\mathcal{F}(\nu)$. We have

$$\begin{aligned}\textstyle\int_E \max\{g, h\}\,d\mu &= \textstyle\int_{E\cap\{g\geq h\}} g\,d\mu + \int_{E\cap\{g<h\}} h\,d\mu \\ &\leq \nu(E \cap \{g \geq h\}) + \nu(E \cap \{g < h\}) = \nu(E),\end{aligned}$$

and hence $\max\{g, h\}$ is indeed in $\mathcal{F}(\nu)$.

With the f_n's now increasing with n, put $f(x) = \lim_n f(x)$. Monotone convergence shows that f is in $\mathcal{F}(\nu)$ and $\int_X f\,d\mu = C$. Define

$$\nu_0(E) = \nu(E) - \int_E f\,d\mu.$$

Then ν_0 is a measure, $\nu_0 \ll \mu$, and the class $\mathcal{F}(\nu_0)$ for ν_0 consists of 0 alone. We shall complete this part of the proof by showing that $\nu_0 = 0$.

If $\nu_0 \neq 0$, choose n large enough so that $\nu_0(X) - \frac{1}{n}\mu(X) > 0$, and put $\nu_0' = \nu_0 - \frac{1}{n}\mu$. Let $X = P \cup N$ be a Hahn decomposition for ν_0' as in Theorem 9.15, and define $g = \frac{1}{n} I_P$. Then the calculation

$$\textstyle\int_E \frac{1}{n} I_P\,d\mu = \frac{1}{n}\mu(P \cap E) = \nu_0(P \cap E) - \nu_0'(P \cap E) \leq \nu_0(P \cap E) \leq \nu_0(E)$$

shows that g is in $\mathcal{F}(\nu_0)$. Hence $g = 0$ a.e. $[d\mu]$, and $\mu(P) = 0$. Since $\nu_0 \ll \mu$, we obtain $\nu_0(P) = 0$ and therefore also $\nu_0'(P) = 0$. Then $\nu_0' \leq 0$, and we must have $\nu_0(X) - \frac{1}{n}\mu(X) \leq 0$. This contradicts the choice of n, and the proof of existence is complete when μ and ν are finite.

PROOF OF EXISTENCE IN THEOREM 9.16 WHEN μ AND ν ARE σ-FINITE. Write X as the countable disjoint union of sets X_n such that $\mu(X_n)$ and $\nu(X_n)$ are both finite. If we put $\mu_n(E) = \mu(E \cap X_n)$ and $\nu_n(E) = \nu(E \cap X_n)$, then μ_n and ν_n are finite measures such that $\nu_n \ll \mu_n$, and the above special case produces functions $f_n \geq 0$ such that $\nu_n(E) = \int_E f_n\,d\mu_n$ for all E. Since $\nu_n(X_n^c) = 0$, we may assume that $f_n(x) = 0$ for $x \notin X_n$. Let $f \geq 0$ be the measurable function that equals f_n on X_n for each n. Then our formula reads $\nu(E \cap X_n) = \int_{E\cap X_n} f\,d\mu$ for all n and for all E. Summing on n, we obtain $\nu(E) = \int_E f\,d\mu$ for all E in $\mathcal{A}$.

Corollary 9.17. Let $(X, \mathcal{A}, \mu)$ be a finite measure space, and let ν be a (real-valued) bounded signed measure on $\mathcal{A}$ with $\nu \ll \mu$ in the sense that $\mu(E) = 0$ implies $\nu(E) = 0$. Then there exists a function f in $L^1(X, \mu)$ such that $\nu(E) = \int_E f\,d\mu$ for all E in $\mathcal{A}$, and f is unique up to a set of μ measure 0.

PROOF. Let $\nu = \nu^+ - \nu^-$ be the Jordan decomposition of ν as in Theorem 9.14, and let $X = P \cup N$ be a Hahn decomposition of ν as in Theorem 9.15. Suppose $\mu(E) = 0$. Since μ is nonnegative, we obtain $\mu(E \cap P) = 0$ and $\mu(E \cap N) = 0$, and the assumption $\nu \ll \mu$ forces

$$0 = \nu(E \cap P) = \nu^+(E \cap P) = \nu^+(E)$$

and $$0 = \nu(E \cap N) = -\nu^-(E \cap N) = -\nu^-(E).$$

Therefore $\nu^+ \ll \mu$ and $\nu^- \ll \mu$, and the corollary follows by applying Theorem 9.16 to ν^+ and ν^- separately.

Corollary 9.18. Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space, and let ν be a σ-finite measure on $\mathcal{A}$. Then there exist a measurable $f \geq 0$ and a set S in $\mathcal{A}$ with $\mu(S) = 0$ such that $\nu = f\,d\mu + \nu_s$, where $\nu_s(E) = \nu(E \cap S)$. The measure ν_s is unique, and the function f is unique up to a set of μ measure 0.

REMARK. The measure ν_s, being carried on a set of μ measure 0, is said to be **singular** with respect to μ. The measure $f\,d\mu$ is, of course, absolutely continuous with respect to μ. The decomposition of ν into the sum of an absolutely continuous part and a singular part is called the **Lebesgue decomposition** of ν with respect to μ. The corollary asserts that this decomposition of measures exists and is unique.

PROOF. As in the proof of Theorem 9.16, we can reduce matters to the case that ν and μ are both finite, and it is therefore enough to handle this special case. Among all sets E in $\mathcal{A}$ with $\mu(E) = 0$, let C be the supremum of $\nu(E)$. The number C is finite, being $\leq \nu(X)$. Choose a sequence of sets E_n in $\mathcal{A}$ with $\mu(E_n) = 0$ and $\nu(E_n)$ increasing to C. Without loss of generality, we may assume that $E_1 \subseteq E_2 \subseteq \cdots$. Put $S = \bigcup_n E_n$. Proposition 5.2 shows that $\mu(S) = 0$ and $\nu(S) = C$. Define $\nu_a(E) = \nu(E \cap S^c)$ and $\nu_s(E) = \nu(E \cap S)$. Then ν_a and ν_s are measures, and $\nu = \nu_a + \nu_s$.

Certainly ν_s is singular with respect to μ, being carried on the set S of μ measure 0. Let us see that ν_a is absolutely continuous. Thus suppose that $\mu(E) = 0$. Then $\mu(S \cup E) \leq \mu(S) + \mu(E) = 0$, and the construction of C shows that $\nu(S \cup E) \leq C = \nu(S)$. Therefore $\nu(S \cup E) - \nu(S) \leq 0$ and $\nu(S \cup E) - \nu(S) = 0$. Hence $0 = \nu(S \cup E) - \nu(S) = \nu(E - S) = \nu(E \cap S^c) = \nu_a(E)$, and ν_a is indeed absolutely continuous. Applying the Radon–Nikodym Theorem (Theorem 9.16), we obtain $\nu = \nu_a + \nu_s = f\,d\mu + \nu_s$. This proves existence.

For uniqueness, suppose that we have $\nu = f\,d\mu + \nu_s = f^\#\,d\mu + \nu_s^\#$ with ν_s and $\nu_s^\#$ carried on respective sets S and $S^\#$ of μ measure 0. The functions f and $f^\#$ are integrable with respect to μ, and we have $\int_E (f - f^\#)\,d\mu = \nu_s^\#(E) - \nu_s(E)$. Taking E to be any subset T in $\mathcal{A}$ of $S \cup S^\#$, we see that $0 = \nu_s^\#(T) - \nu_s(T)$. Therefore $\nu_s^\#(T) = \nu_s(T)$ whenever $T \subseteq S \cup S^\#$. On $(S \cup S^\#)^c$, we have $\nu_s^\#((S \cup S^\#)^c) = \nu_s((S \cup S^\#)^c) = 0$. Therefore $\nu_s^\# = \nu_s$. The uniqueness of the function part follows from the uniqueness in the Radon–Nikodym Theorem, which is part of the statement of that theorem (Theorem 9.16).

5. Continuous Linear Functionals on L^p

We return to the question of identifying the continuous linear functionals on L^p spaces. Let $(X, \mathcal{A}, \mu)$ be a fixed σ-finite measure space. The space $L^p(X, \mu)$ is a normed linear space and, as such, is both a vector space and a metric space. The scalars may be real or complex.

Recall from Section V.9 that a linear functional on $L^p(X, \mu)$ is a linear function from $L^p(X, \mu)$ into the scalars. Proposition 5.57 shows that a linear functional x^* is continuous if and only if it is bounded in the sense that $|x^*(f)| \le C\|f\|_p$ for some constant C and all f in L^p. The inequality $|x^*(f)| \le C\|f\|_p$ holds for all f in L^p if and only if it holds for all f with $\|f\|_p \le 1$, if and only if it holds for all f with $\|f\|_p = 1$. If there is such a constant C, then the finite number

$$\|x^*\| = \sup_{\|f\|_p \le 1} |x^*(f)| = \sup_{\|f\|_p = 1} |x^*(f)|$$

is the least such constant C and is called the **norm** of x^*. Since $\|x^*\|$ is one such constant C, we have

$$|x^*(f)| \le \|x\|^*\|f\|_p.$$

Let p be the dual index to p, defined by $\frac{1}{p} + \frac{1}{p'} = 1$. Each member g of $L^{p'}(X, \mu)$ provides an example of a continuous linear functional on L^p by the formula $x^*(f) = \int_X fg\,d\mu$. The linear functional x^* is bounded, hence continuous, as a consequence of Hölder's inequality: $\left|\int_X fg\,d\mu\right| \le \|g\|_{p'}\|f\|_p$. This inequality shows that $\|x^*\| \le \|g\|_{p'}$. Proposition 9.8 shows that equality $\|x^*\| = \|g\|_{p'}$ holds if μ is σ-finite.

Theorem 9.19 gives a converse when $1 \le p < \infty$, saying that there are no other examples of continuous linear functionals if μ is σ-finite. By contrast, there *can* be other examples in the case of $L^\infty(X, \mu)$. For example, for the situation in which X is the set of positive integers and $\mathcal{A}$ consists of all subsets of X and μ is the counting measure, Problems 39–43 at the end of Chapter V show how to construct a bounded additive set function on $\mathcal{A}$ that is not completely additive, and they show how this set function leads to a notion of integration (hence a linear functional) on this L^∞ space; this linear functional is not given by an L^1 function.

Theorem 9.19 (Riesz Representation Theorem for L^p). Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space, let $1 \leq p < \infty$, and let p' be the dual index to p. If x^* is a continuous linear functional on $L^p(X, \mu)$, then there exists a unique member g of $L^{p'}(X, \mu)$ such that

$$x^*(f) = \int_X fg\, d\mu$$

for all f in L^p. For this function g, $\|x^*\| = \|g\|_{p'}$.

REMARKS. For $1 \leq p < \infty$, Proposition 9.9 shows that $L^p(V, \mu)$ is separable if μ is a Borel measure on an open subset of $\mathbb{R}^N$. For this or any other setting in which any of these L^p spaces is separable, Alaoglu's Theorem (Theorem 5.58) says that any bounded sequence in $L^p(V, \mu)^*$ has a weak-star convergent subsequence. Because of Theorem 9.19 we know what the members of the dual space are. Thus any bounded sequence in $L^{p'}$ has a subsequence that is convergent weak-star against L^p. In effect we obtain a nonconstructive way of producing members of $L^{p'}$. Problem 8 at the end of the chapter will illustrate the usefulness of this technique.

PROOF OF UNIQUENESS. Write $X = \bigcup_{n=1}^\infty X_n$ disjointly with $\mu(X_n)$ finite for all n. If $\int_X fg\, d\mu = 0$ for all f in L^p, then $\int_X I_{A \cap X_n} g\, d\mu = 0$ for every measurable subset A of X. Taking A successively to be each of the sets where $\operatorname{Re} g$ or $\operatorname{Im} g$ is ≥ 0 or is ≤ 0 and applying Corollary 5.23, we see that g is 0 almost everywhere on X_n for each n. Hence g is 0 almost everywhere.

PROOF OF EXISTENCE IF $\mu(X)$ IS FINITE. Temporarily let us suppose that the underlying scalars are real. Define a set function ν on $\mathcal{A}$ by $\nu(E) = x^*(I_E)$; ν is well defined because every I_E is in L^p, and ν is additive because x^* is linear. If E_n is an increasing sequence of measurable sets with union E, then $\lim_n I_{E_n} = I_E$ pointwise, and hence $\lim_n |I_E - I_{E_n}|^p = 0$ pointwise. By dominated convergence, $\lim_n \|I_E - I_{E_n}\|_p = 0$. Thus

$$|\nu(E) - \nu(E_n)| = |x^*(I_E - I_{E_n})| \leq \|x^*\| \|I_E - I_{E_n}\|_p,$$

and the right side has limit 0. By Proposition 5.2, ν is completely additive. The set function ν is bounded because $|\nu(E)| = |x^*(I_E)| \leq \|x^*\| \|I_E\|_p = \|x^*\|(\mu(E))^{1/p} \leq \|x^*\|(\mu(X))^{1/p}$, and it satisfies $\nu \ll \mu$ because if $\mu(E) = 0$, then I_E is the 0 function of L^p and thus $\nu(E) = x^*(I_E) = x^*(0) = 0$. By the Radon–Nikodym Theorem in the form of Corollary 9.17, there exists an integrable real-valued function g such that $\nu(E) = \int_E g\, d\mu$ for all E, i.e.,

$$x^*(I_E) = \int_X I_E g\, d\mu \qquad \text{for every measurable set } E.$$

By linearity, this equality extends to show that $x^*(s) = \int_X sg\,d\mu$ for every simple function s. Let $f \geq 0$ be in L^p, and choose an increasing sequence $\{s_n\}$ of simple functions ≥ 0 with pointwise limit f. We shall show that fg is integrable and $x^*(f) = \int fg\,d\mu$. In fact, let A be the set where $g(x) \geq 0$. Then $\lim_n |fI_A - s_nI_A|^p = 0$ pointwise, and hence $\lim_n \|fI_A - s_nI_A\|_p = 0$ by dominated convergence. Since

$$|x^*(fI_A) - x^*(s_nI_A)| \leq \|x^*\|\|fI_A - s_nI_A\|_p$$

and since the right side tends to 0, the set $\{x^*(s_nI_A)\}$ of numbers is bounded. Thus the set $\left\{\int_X s_nI_Ag\,d\mu\right\}$ of equal numbers is bounded. Since $g \geq 0$ on A, the functions s_nI_Ag increase to fI_Ag, and thus $\int_X fI_Ag\,d\mu$ is finite by monotone convergence. In other words, fg^+ is integrable. Similarly fg^- is integrable, and thus fg is integrable. Since $\lim_n x^*(s_nI_A) = x^*(fI_A)$ and $\lim_n \int_X s_nI_Ag\,d\mu = \int_X fI_Ag\,d\mu$ and since a similar result holds for g^-, we conclude that

$$x^*(f) = \int_X fg\,d\mu \qquad \text{for all } f \geq 0 \text{ in } L^p.$$

This conclusion, now proved for $f \geq 0$, immediately extends by linearity to all f in L^p and completes the verification that $x^*(f) = \int_X fg\,d\mu$ in the case that the scalars are real.

If the scalars are complex, we apply the above argument to the restrictions of $\operatorname{Re} x^*$ and $\operatorname{Im} x^*$ to the real-valued functions in L^p, obtaining real-valued functions g_1 and g_2 in $L^{p'}$ with $\operatorname{Re} x^*(f) = \int_X fg_1\,d\mu$ and $\operatorname{Im} x^*(f) = \int_X fg_2\,d\mu$ for all real-valued f. Then $x^*(f) = \int_X f(g_1 + ig_2)\,d\mu$ for all real-valued f, and it follows that this same equality is valid for all complex-valued f. Since g_1 and g_2 are in L^p, so is $g_1 + ig_2$. This completes the verification that $x^*(f) = \int_X fg\,d\mu$ for a suitable g in the case that the scalars are complex.

Finally Proposition 9.8 shows that $\|x^*\| = \|g\|_{p'}$ and completes the proof of the theorem under the assumption that $\mu(X)$ is finite.

PROOF OF EXISTENCE IF $\mu(X)$ IS σ-FINITE. Again we temporarily suppose that the underlying scalars are real. Since μ is σ-finite, we can write X as the increasing union of sets E_n of finite measure. Let L_n^p be the set of members of L^p that vanish off E_n, and let x_n^* be the restriction of x^* to L_n^p. Find, by the special case just completed, a function g_n for each n such that $x_n^*(f_n) = \int_{E_n} f_ng_n\,d\mu$ for all f_n in L_n^p. The already proved uniqueness result implies that the restriction of g_{n+1} to E_n equals g_n almost everywhere $[d\mu]$. Let g be the measurable function equal to g_1 on E_1 and equal to g_n on $E_n - E_{n-1}$ if $n \geq 2$. Let A be the set where $g(x) \geq 0$, and let $f \geq 0$ be in L^p. Then $fI_{E_n \cap A}$ increases to fI_A, and dominated convergence implies that $\lim_n \|fI_{E_n\cap A} - fI_A\|_p = 0$. Since $fI_{E_n\cap A}g$ increases pointwise to fI_Ag, monotone convergence gives

$$\begin{aligned}\int_X fI_Ag\,d\mu &= \lim_n \int_X fI_{E_n\cap A}g\,d\mu = \lim_n \int_X fI_{E_n\cap A}g_n\,d\mu \\ &= \lim_n x_n^*(fI_{E_n\cap A}) = \lim_n x^*(fI_{E_n\cap A}) = x^*(fI_A),\end{aligned}$$

the last equality holding since $\|fI_{E_n\cap A} - fI_A\|_p$ tends to 0. Hence fg^+ is integrable. By proceeding similarly with the set where $g(x) < 0$ and by writing a general f as $f = f^+ - f^-$, we conclude that fg is integrable for every f in L^p and $x^*(f) = \int_X fg\,d\mu$, provided the scalars are real.

Again there is no difficulty in extending the argument to the case that the scalars are complex, and Proposition 9.8 shows that $\|x^*\| = \|g\|_{p'}$.

6. Marcinkiewicz Interpolation Theorem

This section concerns linear functions and some almost-linear functions between L^p spaces. We saw evidence in Proposition 9.5 that the L^p spaces behave collectively like a well-behaved family of spaces. That result specifically gave an upper bound for $\|f\|_p$ in terms of $\|f\|_{p_1}$ and $\|f\|_{p_2}$ when $p_1 \le p \le p_2$. It turns out that linear functions between pairs of L^p spaces satisfy inequalities of a similar sort.

There are two classes of results in this direction. Results of the first kind use methods of complex analysis, address bounded linear operators only, and give estimates for a one-parameter family of operators that are sharp at the ends. The main result of this kind is the "Riesz[1] Convexity Theorem," whose precise general statement and proof we omit. The thrust of the theorem is that if a linear operator T satisfies the two estimates $\|T(f)\|_{q_1} \le M_1\|f\|_{p_1}$ and $\|T(f)\|_{q_2} \le M_2\|f\|_{p_2}$, then T satisfies also an estimate $\|T(f)\|_q \le M\|f\|_p$ for all pairs (p, q) such that $\left(\frac{1}{p}, \frac{1}{q}\right)$ lies on the line segment in the $\left(\frac{1}{p}, \frac{1}{q}\right)$ plane from $\left(\frac{1}{p_1}, \frac{1}{q_1}\right)$ to $\left(\frac{1}{p_2}, \frac{1}{q_2}\right)$. The conclusion gives also some specific information about M.

The existence of some M in the Riesz theorem can be obtained in most cases of interest by a corresponding real-analysis result known as the "Marcinkiewicz Interpolation Theorem." We include below a statement of the Marcinkiewicz theorem in general and the proof in a special case of exceptional interest. The Marcinkiewicz theorem imposes some restrictions on the pairs (p_1, q_1) and (p_2, q_2) that are not needed in the Riesz theorem, but situations that do not

[1]The person in question here is Marcel Riesz, whose name is associated also with convergence of the partial sums of the Fourier series of an L^p function in L^p for $1 < p < \infty$. The other mentions of the name "Riesz" in this book, namely in connection with the Rising Sun Lemma of Section VII.1 and various results known as the Riesz Representation Theorem, refer to Frigyes Riesz.

satisfy these restrictions are of comparatively little interest in applications. In any event, in the situations where the Marcinkiewicz theorem applies, it is only the specific information about M in the Riesz theorem that does not come out of the real-analysis proof of the Marcinkiewicz theorem.

Let us mention without proof two consequences of the Riesz Convexity Theorem—the Hausdorff–Young Theorem and Young's inequality.

The linear operator T in the Hausdorff–Young Theorem is the Fourier transform $\mathcal{F}$, and the instances of the theorem that we knew previously are when (p, p') equals $(1, \infty)$ or $(2, 2)$. The numerology that allows the Riesz Convexity Theorem to apply is that

$$\frac{1}{p} = \frac{1-t}{1} + \frac{t}{2} \qquad \text{and} \qquad \frac{1}{p'} = \frac{1-t}{\infty} + \frac{t}{2}$$

for the same t:

HAUSDORFF–YOUNG THEOREM. If $1 \leq p \leq 2$ and if p' is the dual index, then the Fourier transform $\mathcal{F}$, initially defined on the dense subspace $L^1(\mathbb{R}^N) \cap L^2(\mathbb{R}^N)$ of $L^p(\mathbb{R}^N)$, satisfies

$$\|\mathcal{F}(f)\|_{p'} \leq \|f\|_p$$

for such f and therefore extends to all of $L^p(\mathbb{R}^N)$ in such a way that this same inequality holds.

If one tries to derive the Hausdorff–Young Theorem from the Marcinkiewicz Interpolation Theorem, one gets only the conclusion $\|\mathcal{F}(f)\|_{p'} \leq M\|f\|_p$ without the improvement on the bound: $M \leq 1$.

The linear operator T in Young's inequality can be taken to be $g \mapsto f * g$ with f fixed in L^p. The instances of the inequality that we knew previously are when (q, r) equals $(1, p)$ or (p', ∞). The relevant numerology is that

$$\frac{1}{q} = \frac{1-t}{1} + \frac{t}{p'} \qquad \text{and} \qquad \frac{1}{r} = \frac{1-t}{p'} + \frac{t}{\infty}$$

for the same t:

YOUNG'S INEQUALITY. Let p, q, and r be three indices ≥ 1 and $\leq \infty$ such that $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$. Then convolution $f * g$ is well defined for f in $L^p(\mathbb{R}^N)$ and g in $L^q(\mathbb{R}^N)$, and it satisfies

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

By way of preparation for the statement of the Marcinkiewicz theorem, let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let T be a function from a vector subspace of measurable functions on X, modulo sets[2] of μ measure 0, into measurable functions on Y, modulo sets of ν measure 0. We say that T is a **sublinear** operator if $|T(f+g)| \le |T(f)| + |T(g)|$ for all f and g in the domain of T.

The two examples of T to keep in mind are the sublinear operator $f \mapsto f^*$ in $\mathbb{R}^N$ of passing to the Hardy–Littlewood maximal function, as in Section VI.6, and the linear operator $f \mapsto H_1 f$ in $\mathbb{R}^1$ of forming a certain approximation H_1 to the Hilbert transform, as in Section VIII.7. More specifically the Hardy–Littlewood maximal function of a locally integrable function f on $\mathbb{R}^N$ is defined as

$$f^*(x) = \sup_{0<r<\infty} m(B_r)^{-1} \int_{B_r} |f(x-y)|\,dy, \qquad \text{where } B_r = B(r;0) \text{ in } \mathbb{R}^N,$$

and the sublinear operator T is $Tf = f^*$. The approximation H_1 to the Hilbert transform is defined for f in $L^1 + L^2$ by

$$H_1 f(x) = h_1 * f(x) = \frac{1}{\pi} \int_{|t|\ge 1} \frac{f(x-t)}{t}\,dt$$

as the convolution with a fixed L^2 function.

Let $1 \le p, q \le \infty$. We generalize the notion of boundedness of a linear operator between $L^p(X)$ and $L^q(Y)$ so that we can work with sublinear operators as well as linear ones. A sublinear operator T is said to be of **type** (p,q) or **strong type** (p,q) if $\|Tf\|_q \le M\|f\|_p$ with M finite and independent of f. The least M for which this inequality holds is called the **norm** or **operator norm** of T. If $q < \infty$, then Chebyshev's inequality from Section VI.10 gives

$$\nu\big(\{y \in Y \mid |Tf(y)| > \xi\}\big) \le \frac{\int_Y |Tf|^q\,d\nu}{\xi^q},$$

and for any M such that $\|Tf\|_q \le M\|f\|_p$ for all f, it follows that

$$\nu\big(\{y \in Y \mid |Tf(y)| > \xi\}\big) \le \left(\frac{M\|f\|_p}{\xi}\right)^q.$$

If $q < \infty$, a sublinear operator T is said to be of **weak type** (p,q) if it satisfies

$$\nu\big(\{y \in Y \mid |Tf(y)| > \xi\}\big) \le \left(\frac{M\|f\|_p}{\xi}\right)^q$$

[2] This condition means that the domain of T is to be regarded as a vector subspace of measurable functions, except that two functions are identified if they differ only on a set of measure 0.

for some M. In this case the least such M is called the **weak-type norm** of T. We already encountered the definition of weak type $(1, 1)$ in Section VI.6. If $q = \infty$, the convention is that weak type (p, ∞) is the same as strong type (p, ∞).

Consider our two examples. The operation $T(f) = f^*$ of passing to the Hardy–Littlewood maximal function in $\mathbb{R}^N$ is of weak type $(1, 1)$ by the Hardy–Littlewood Maximal Theorem (Theorem 6.38), and the evident inequality

$$\Big\| \sup_{0<r<\infty} m(B_r)^{-1} \int_{B_r} |f(x-y)|\, dy \Big\|_\infty \leq \|f\|_\infty$$

shows that $f \mapsto f^*$ is of type (∞, ∞) as well. The linear operator $T(f) = H_1 f$ of passing to the approximation H_1 to the Hilbert transform in $\mathbb{R}^1$ is of weak type $(1, 1)$ and type $(2, 2)$ by Theorem 8.25.

Theorem 9.20 (Marcinkiewicz Interpolation Theorem). Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let (p_1, q_1) and (p_2, q_2) be two pairs of indices between 1 and ∞. Suppose that $1 \leq p_1 \leq q_1 \leq \infty$, $1 \leq p_2 \leq q_2 \leq \infty$, and $p_1 \neq p_2$. Let T be a sublinear operator from $L^{p_1}(X, \mu) + L^{p_2}(X, \mu)$ to the space of measurable functions on Y modulo sets of ν measure 0, and suppose that T is of weak types (p_1, q_1) and (p_2, q_2) with respective weak-type norms M_1 and M_2. Fix t with $0 < t < 1$, and define (p, q) by

$$\frac{1}{p} = \frac{1-t}{p_1} + \frac{t}{p_2} \qquad \text{and} \qquad \frac{1}{q} = \frac{1-t}{q_1} + \frac{t}{q_2}.$$

Then T is of strong type (p, q) with

$$\|Tf\|_q \leq C\|f\|_p \qquad \text{for all } f \in L^p(X, \mu),$$

with the constant C depending only on $t, M_1, M_2, p_1, q_1, p_2, q_2$ and with C bounded as a function of t as long as t is bounded away from 0 and 1.

Before discussing the proof, let us apply the theorem to our two examples, the Hardy–Littlewood maximal function and the approximation H_1 to the Hilbert transform. Then let us draw some consequences of these applications. As was said before the statement of Theorem 9.20, the sublinear operator $f \mapsto f^*$ is of weak type $(1, 1)$ and strong type $(2, 2)$. The theorem immediately gives the following corollary.

Corollary 9.21. If $1 < p \leq \infty$, then there exists a constant A_p such that the Hardy–Littlewood maximal function satisfies

$$\|f^*\|_p \leq A_p \|f\|_p$$

for all f in $L^p(\mathbb{R}^N)$.

The case of this result in one dimension implies something in N dimensions that we have not obtained earlier. If f is locally integrable on $\mathbb{R}^N$, one says that **strong differentiation** holds for f at x if

$$\lim_{\substack{\operatorname{diam}(R)\to 0,\\ R=\text{geometric rectangle}\\ \text{centered at } x}} \frac{1}{m(R)} \int_R f(y)\,dy = f(x).$$

A consequence of Corollary 9.21 is that strong differentiation holds almost everywhere for each f in $L^p(\mathbb{R}^N)$ for $p > 1$. The proof is outlined in Problems 13–15 at the end of the chapter. By contrast, it is known that there are functions in $L^1(\mathbb{R}^N)$ for which strong differentiation fails everywhere.

In the second example the operator H_1 that approximates the Hilbert transform is of weak type $(1, 1)$ and strong type $(2, 2)$, and Theorem 9.20 allows us to conclude that it is of strong type (p, p) for $1 < p \le 2$. But we can do better. The operator H_1 is convolution by the function h_1 with $h_1(x) = 1/(\pi x)$ for $|x| \ge 1$ and $h_1(x) = 0$ for $|x| < 1$. The function h_1 is in L^p for all $p > 1$, and Proposition 9.10f shows that $h_1 * f$ is well defined as a bounded continuous function whenever f is in some L^q with $1 \le q < \infty$. Thus H_1 is defined on all L^p classes for $1 < p < \infty$, and a general result that we prove below as Lemma 9.22 shows that an inequality $\|H_1 f\|_p \le A_p \|f\|_p$ for all f in L^p implies $\|H_1 g\|_{p'} \le A_p \|g\|_{p'}$ for all g in $L^{p'}$, provided p' is the dual index to p and $1 < p < \infty$. Thus the boundedness result for H_1 on L^p extends to $1 < p < \infty$.

Next, we define the dilate $h_\varepsilon(x) = \varepsilon^{-1} h_1(x)$ in the usual way and put $H_\varepsilon f = h_\varepsilon * f$. We shall see for every $\varepsilon > 0$ that $\|H_\varepsilon f\|_p \le A_p \|f\|_p$ with the same constant A_p, and finally we shall see that we can let ε decrease to 0 and obtain the Hilbert transform H as a well-defined linear operator on all L^p classes for $1 < p < \infty$; the estimate is $\|Hf\|_p \le A_p \|f\|_p$, again for the same A_p. Problems 20–22 at the end of the chapter indicate how to use this boundedness to prove that the Fourier series of any L^p function on $[-\pi, \pi]$ converges to the function in L^p if $1 < p < \infty$.

Lemma 9.22. Fix p with $1 < p < \infty$, let p' be the dual index, and suppose that h is in $L^p(\mathbb{R}^N) \cap L^{p'}(\mathbb{R}^N)$. If $\|h * f\|_p \le A_p \|f\|_p$ for all f in $L^p(\mathbb{R}^N)$, then $\|h * g\|_{p'} \le A_p \|g\|_{p'}$ for all g in $L^{p'}$.

REMARKS. Since h is in $L^{p'}$, $h * f$ is in L^∞ when f is in L^p. Thus $h * f$ is well defined, and it is meaningful to say that $h * f$ is actually in L^p. When $h * f$ is in L^p, the integral $\int (h * f) g\,dx$ is well defined for g in $L^{p'}$. A little care is required in working with this integral in the proof because $\int (|h| * f) g\,dx$ need not be well defined and Fubini's Theorem may not directly applicable.

PROOF. For any function F on $\mathbb{R}^N$, define $F^\#(x) = F(-x)$ and observe that $\|F^\#\|_r = \|F\|_r$ for $1 \le r \le \infty$. If g is an integrable simple function, then $(h^\# * g)(x) = \int h(y-x)g(y)\,dy = \int h(-y-x)g^\#(y)\,dy = (h * g^\#)(-x)$. Thus this g and an integrable simple function f together satisfy

$$\begin{aligned}\int (h * f^\#)(x)g(x)\,dx &= \iint h(x-y)f(-y)g(x)\,dy\,dx \\ &= \iint h(x+y)f(y)g(x)\,dy\,dx\end{aligned}$$

and
$$\begin{aligned}\int (h * g^\#)(y)f(y)\,dy &= \int (h^\# * g)(-y)f(y)\,dy \\ &= \iint h^\#(-y-x)g(x)f(y)\,dx\,dy \\ &= \iint h(x+y)g(x)f(y)\,dx\,dy.\end{aligned}$$

Because f and g are in every L^r class, the right sides of these two displays are finite when absolute value signs are inserted in the integrands. Thus Fubini's Theorem applies and shows that the two right sides are equal. Combining this fact with Hölder's inequality and the hypothesis about h, we obtain

$$\begin{aligned}\Big|\int (h * g^\#)(y)f(y)\,dy\Big| &= \Big|\int (h * f^\#)(x)g(x)\,dx\Big| \\ &\le \|h * f^\#\|_p\|g\|_{p'} \le A_p\|f^\#\|_p\|g\|_{p'} = A_p\|f\|_p\|g\|_{p'}\end{aligned}$$

whenever f and g are integrable simple functions. If a general f_0 in L^p is given, we can find a sequence f_n of integrable simple functions such that $\|f_n - f_0\|_p \to 0$, and we apply this inequality to each f_n. Then the left side of the inequality tends to $\big|\int (h * g^\#)(y)f_0(y)\,dy\big|$, and the right side tends to $A_p\|f_0\|_p\|g\|_{p'}$. Taking the supremum over all f_0 with $\|f_0\|_p \le 1$ and applying Proposition 9.8, we find that $\|h * g^\#\|_{p'} \le A_p\|g\|_{p'} = A_p\|g^\#\|_{p'}$. In other words,

$$\|h * g_n\|_{p'} \le A_p\|g_n\|_{p'}$$

for every integrable simple function g_n. For a general g in $L^{p'}$, choose a sequence of integrable simple functions g_n with $\|g_n - g\|_{p'} \to 0$. Since h is in L^p, it follows from Proposition 9.10f that $h * g_n$ converges to $h * g$ uniformly. On the other hand, the inequality $\|h * (g_m - g_n)\|_{p'} \le A_p\|g_m - g_n\|_{p'}$ shows that $\{h * g_n\}$ is Cauchy in $L^{p'}$. By Theorem 5.58, $\{h * g_n\}$ converges to some function in $L^{p'}$ and has an almost-everywhere convergent subsequence to this function. Since $h * g_n$ converges uniformly to $h * g$, we conclude that $h * g_n$ converges to $h * g$ in $L^{p'}$. Therefore $\|h * g\|_{p'} \le A_p\|g\|_{p'}$, and the proof is complete.

Again let h_1 be the function on $\mathbb{R}^1$ equal to $1/(\pi x)$ for $|x| \ge 1$ and equal to 0 for $|x| < 1$. This is in $L^r(\mathbb{R}^1)$ for every $r > 1$. Our operator giving an approximation to the Hilbert transform is $H_1 f = h_1 * f$. Using our results from Chapter VIII along with the Marcinkiewicz Interpolation Theorem, we saw earlier in this section that H_1 satisfies $\|H_1 f\|_p \le A_p \|f\|_p$ for $1 < p \le 2$ and all f in $L^p(\mathbb{R}^1)$. Lemma 9.22 shows that this inequality remains valid for $1 < p < \infty$. From this result we can extend the Hilbert transform to $L^p(\mathbb{R}^1)$ for all p with $1 < p < \infty$, as follows.

Theorem 9.23. Let $1 < p < \infty$, let

$$h_\varepsilon(x) = \varepsilon^{-1} h_1(\varepsilon^{-1}x) = \begin{cases} 1/(\pi x) & \text{for } |x| \ge \varepsilon, \\ 0 & \text{for } |x| < \varepsilon, \end{cases}$$

and define $H_\varepsilon f = h_\varepsilon * f$ for f in L^p and $\varepsilon > 0$. Then

(a) there exists a constant A_p independent of ε such that $\|H_\varepsilon f\|_p \le A_p \|f\|_p$ for all f in L^p,

(b) the limit
$$Hf(x) = \lim_{\varepsilon \downarrow 0} \frac{1}{\pi} \int_{|t| \ge \varepsilon} \frac{f(x-t)\,dt}{t}$$
exists in L^p for every f in L^p,

(c) the operator H satisfies $\|Hf\|_p \le A_p \|f\|_p$ for every f in L^p.

PROOF. Convolution with h_ε is well defined on L^p because h_ε is in $L^{p'}$, p' being the dual index for p. The three computations

$$\begin{aligned} H_\varepsilon f(x) = (f * h_\varepsilon)(x) &= \textstyle\int f(x-y)\varepsilon^{-1} h_1(\varepsilon^{-1}y)\,dy = \int f(x - \varepsilon y) h_1(y)\,dy \\ &= \textstyle\int \varepsilon^{-1} f_{\varepsilon^{-1}}(\varepsilon^{-1}x - y) h_1(y)\,dy = \varepsilon^{-1}(H_1 f_{\varepsilon^{-1}})(\varepsilon^{-1}x), \end{aligned}$$

$$\textstyle\int |(H_\varepsilon f)(x)|^p\,dx = \varepsilon^{-p} \int |(H_1 f_{\varepsilon^{-1}})(\varepsilon^{-1}x)|^p\,dx = \varepsilon^{1-p} \int |(H_1 f_{\varepsilon^{-1}})(x)|^p\,dx,$$

and $$\textstyle\int |g_{\varepsilon^{-1}}(x)|^p\,dx = \varepsilon^p \int |g(\varepsilon x)|^p\,dx = \varepsilon^{-1+p} \int |g(x)|^p\,dx$$

allow us to write

$$\|H_\varepsilon f\|_p^p = \varepsilon^{1-p} \|H_1 f_{\varepsilon^{-1}}\|_p^p \le A_p^p\, \varepsilon^{1-p} \|f_{\varepsilon^{-1}}\|_p^p = A_p^p \|f\|_p^p.$$

This proves (a), the constant A_p being any constant that works for H_1.

In Lemma 9.24 below we show by a direct computation that (b) holds for the dense subset of C^1 functions f of compact support. Let us deduce (b) for general f in L^p from this fact and (a). In fact, if we are given f, we choose a sequence f_n in the dense set with $f_n \to f$ in L^p. Then

$$\begin{aligned} \|H_\varepsilon f - H_{\varepsilon'} f\|_p &\le \|H_\varepsilon (f - f_n)\|_p + \|H_\varepsilon f_n - H_{\varepsilon'} f_n\|_p + \|H_{\varepsilon'}(f_n - f)\|_p \\ &\le A_p \|f_n - f\|_p + \|H_\varepsilon f_n - H_{\varepsilon'} f_n\|_p + A_p \|f_n - f\|_p. \end{aligned}$$

Choose n to make the first and third terms small on the right, and then choose ε and ε' sufficiently close to 0 so that the second term on the right is small. The result is that $H_{\varepsilon_n} f$ is Cauchy in L^p along any sequence ε_n tending to 0. This proves (b), apart from the direct computation for the dense subset.

In (b), we proved that $H_\varepsilon f \to Hf$ in L^p. Then (a) gives $\|Hf\|_p = \lim_{\varepsilon\downarrow 0} \|H_\varepsilon f\|_p \leq \limsup_{\varepsilon\downarrow 0} A_p\|f\|_p = A_p\|f\|_p$. This proves (c) and completes the proof of Theorem 9.23 except for the following lemma.

Lemma 9.24. If f is a C^1 function of compact support on $\mathbb{R}^1$, then

$$\lim_{\varepsilon\downarrow 0} \frac{1}{\pi} \int_{|t|\geq\varepsilon} \frac{f(x-t)\,dt}{t}$$

exists uniformly and in L^p for every $p > 1$.

PROOF. Let $\|\cdot\|$ denote the supremum norm or the L^p norm. By the Cauchy criterion it is enough to show that

$$\Big\| \int_{\varepsilon_1\leq|t|\leq\varepsilon_2} \frac{f(x-t)\,dt}{t} \Big\|$$

tends to 0 for the above interpretations of $\|\cdot\|$ as ε_1 and ε_2 tend to 0. Since $|f'(u)| \leq M$, use of the Mean Value Theorem on $\operatorname{Re} f$ and $\operatorname{Im} f$ shows that $|f(x-t)-f(x)| \leq 2M|t|$. Suppose that $0 < \varepsilon_1 \leq \varepsilon_2 \leq 1$. If E is a compact set containing the sum of any member of the support of f and any x with $|x| \leq 1$, then it follows that

$$\begin{aligned}
\Big\| \int_{\varepsilon_1\leq|t|\leq\varepsilon_2} \frac{f(x-t)\,dt}{t} \Big\| &= \Big\| \int_{\varepsilon_1\leq|t|\leq\varepsilon_2} \frac{[f(x-t)-f(x)]\,dt}{t} \Big\| \\
&\leq \int_{\varepsilon_1\leq|t|\leq\varepsilon_2} \frac{\|f(x-t)-f(x)\|_x\,dt}{|t|} \\
&\leq \int_{\varepsilon_1\leq|t|\leq\varepsilon_2} \frac{2M|t|\|I_E\|\,dt}{|t|} \\
&= 4M\|I_E\|(\varepsilon_2-\varepsilon_1).
\end{aligned}$$

The right side tends to 0 as ε_1 and ε_2 tend to 0, and the proof of the lemma is complete.

Having now completely proved Theorem 9.23, let us return to a discussion of the proof of the Marcinkiewicz theorem, Theorem 9.20. The proof is considerably simplified by assuming that $q_1 = p_1$ and $q_2 = p_2$, which happens to be the special case of most interest to us, and we shall give a proof only under this additional hypothesis. The idea in the special case will be to estimate integrals of powers of functions by using Proposition 6.56b to reduce the estimates to facts about distribution functions.

The proof in general has the same flavor as the argument we give, but it involves also a subtler decomposition of f into two parts, a nonobvious application of Hölder's inequality, and a clever use of Proposition 9.8.

PROOF OF THEOREM 9.20 WHEN $p_1 = q_1 < p_2 = q_2$. We divide matters into two cases, the first when $p_2 < \infty$ and the second when $p_2 = \infty$.

We begin with the case with $p_2 < \infty$. Let

$$\lambda(\xi) = \lambda_{Tf}(\xi) = \nu\big(\big\{y \bigm| |Tf(y)| > \xi\big\}$$

be the distribution function of Tf as in Section VI.10. Proposition 6.56b shows that

$$\|Tf\|_p^p = p\int_0^\infty \xi^{p-1}\lambda(\xi)\,d\xi = 2^p p\int_0^\infty \xi^{p-1}\lambda(2\xi)\,d\xi. \qquad (*)$$

With $\xi > 0$ fixed, we shall estimate $\lambda(2\xi)$. We decompose f as $f = f_1 + f_2$ with

$$f_1(x) = \left\{\begin{matrix} f(x) & \text{if } |f(x)| > \xi \\ 0 & \text{otherwise}\end{matrix}\right\} \qquad\text{and}\qquad f_2(x) = \left\{\begin{matrix} f(x) & \text{if } |f(x)| \le \xi \\ 0 & \text{otherwise}\end{matrix}\right\}.$$

Just as in the proof of Proposition 9.4c, f_1 is in $L^{p_1}(X,\mu)$ and f_2 is in $L^{p_2}(X,\mu)$. Because $f = f_1 + f_2$, sublinearity of T gives $|Tf| \le |Tf_1| + |Tf_2|$. If λ_1 and λ_2 are the distribution functions of Tf_1 and Tf_2 and if $\alpha > 0$ is given, then

$$\lambda(2\alpha) \le \lambda_1(\alpha) + \lambda_2(\alpha)$$

because $|Tf|$ can be $> 2\alpha$ only if at least one of $|Tf_1|$ and $|Tf_2|$ is $> \alpha$. For every $\alpha > 0$, the assumption that T is of weak types (p_1, p_1) and (p_2, p_2) gives us

$$\lambda_1(\alpha) \le \left(\frac{M_1\|f_1\|_{p_1}}{\alpha}\right)^{p_1} \qquad\text{and}\qquad \lambda_2(\alpha) \le \left(\frac{M_2\|f_2\|_{p_2}}{\alpha}\right)^{p_2}.$$

For $\alpha = \xi$, we therefore obtain

$$\begin{aligned}\lambda(2\xi) \le \lambda_1(\xi) + \lambda_2(\xi) &\le M_1^{p_1}\xi^{-p_1}\int_X |f_1|^{p_1}\,d\mu + M_2^{p_2}\xi^{-p_2}\int_X |f_2|^{p_2}\,d\mu \\ &= M_1^{p_1}\xi^{-p_1}\int_{\{|f|>\xi\}} |f|^{p_1}\,d\mu + M_2^{p_2}\xi^{-p_2}\int_{\{|f|\le\xi\}} |f|^{p_2}\,d\mu. \qquad (**)\end{aligned}$$

With the estimate for $\lambda(2\xi)$ in hand, we can now let ξ vary and estimate $\|Tf\|_p^p$. From $(*)$ and $(**)$ we obtain $\|Tf\|_p^p \le I_1 + I_2$, where

$$I_1 = 2^p p M_1^{p_1} \int_0^\infty \xi^{p-p_1-1} \int_{\{|f(x)|>\xi\}} |f(x)|^{p_1}\, d\mu(x)\, d\xi$$

and
$$I_2 = 2^p p M_2^{p_2} \int_0^\infty \xi^{p-p_2-1} \int_{\{|f(x)|\le\xi\}} |f(x)|^{p_2}\, d\mu(x)\, d\xi.$$

Fubini's Theorem gives

$$I_1 = 2^p p M_1^{p_1} \int_X |f|^{p_1} \Big[\int_0^{|f|} \xi^{p-p_1-1}\, d\xi \Big]\, d\mu = \frac{2^p p M_1^{p_1}}{p-p_1} \int_X |f|^p\, d\mu.$$

Similarly

$$I_2 = \frac{2^p p M_2^{p_2}}{p_2 - p} \int_X |f|^p\, d\mu,$$

and thus $\|Tf\|_p^p \le C^p \|f\|_p^p$ as required.

The remaining case to handle has $p_2 = \infty$. The general line of the argument is the same as above, but there are small differences. With ξ fixed, the definitions of f_1 and f_2 are adjusted to be

$$f_1(x) = \begin{cases} f(x) & \text{if } |f(x)| > \xi/\|T\|_\infty, \\ 0 & \text{otherwise,} \end{cases}$$

and $f_2 = f - f_1$. Then $\|f_2\|_\infty \le \xi \big/ \|T\|_\infty$, $\|Tf_2\|_\infty \le \xi$, and $\lambda_2(\xi) = 0$. Hence

$$\lambda(2\xi) \le \lambda_1(\xi) + \lambda_2(\xi) = \lambda_1(\xi) \le M_1^{p_1} \xi^{-p_1} \int_{\{|f|>\xi/\|T\|_\infty\}} |f|^{p_1}\, d\mu,$$

and then the proof can proceed along the lines above.

7. Problems

1. For a measure space of finite measure, prove that $L^p \subseteq L^q$ whenever $p \ge q \ge 1$. More particularly prove, for the case that the total measure is 1, that $\|f\|_q \le \|f\|_p$ whenever $p \ge q \ge 1$.
2. Let p, q, r be real numbers in $[1, +\infty]$ with $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$. Using the equality $\frac{r'}{p} + \frac{r'}{q} = 1$ and Hölder's inequality, prove that $\int_X |fgh|\, d\mu \le \|f\|_p \|g\|_q \|h\|_r$.

3. For a measure space of finite measure, let $\{f_n\}$ be a sequence of measurable functions converging pointwise to f. Suppose that $1 \le q < p < \infty$, and suppose that the sequence of numbers $\{\|f\|_p\}$ is bounded. Using Egoroff's Theorem (Problem 17, Chapter V) or uniform integrability (Problem 21, Chapter V), prove that $f_n \to f$ in L^q.

4. This problem produces an example of a measure space in which two distinct members of L^∞ act as the same linear functional on L^1. The measure space $(X, \mathcal{A}, \mu)$ has X consisting of a single point p, $\mathcal{A} = \{\varnothing, X\}$, and $\mu(X) = +\infty$.
   (a) Show that $\dim L^1(X) = 0$ and $\dim L^\infty(X) = 1$.
   (b) Proposition 9.8 assumed σ-finiteness to ensure its conclusion when $p = \infty$. Show that the conclusion of Proposition 9.8 fails for $p = \infty$ in this example.

5. If f is real-valued and integrable on the measure space $(X, \mathcal{A}, \mu)$, what are all the Hahn decompositions for the signed measure $\nu(E) = \int_E f\,d\mu$?

6. Provide examples of each of the following. Each example can be produced on one of the following three algebras of subsets of a set X: the finite subsets of a X and their complements, all subsets of a countable set X, the Borel sets of $X = [0, 1]$.
   (a) An additive set function ν on an algebra of sets with $|\nu(X)| < \infty$ but with $\sup_E |\nu(E)| = \infty$.
   (b) A counterexample to the Hahn decomposition if the assumption "σ-algebra" is relaxed to "algebra" but the other assumptions are left in place.
   (c) A finite measure ν and a non σ-finite measure μ, both defined on a σ-algebra, such that $\nu \ll \mu$ but ν is not given by an integral with respect to μ.

Problems 7–8 concern harmonic functions and the Poisson integral formula for the unit disk in $\mathbb{R}^2$. These matters were the subject of Problems 27–29 at the end of Chapter I, Problems 14–15 at the end of Chapter III, Problems 10–13 at the end of Chapter IV, and Problems 18–20 at the end of Chapter VI. Problem 7 updates the results from Chapter VI so that they apply for $1 \le p < \infty$, and Problem 8 uses weak-star convergence to establish a converse result.

7. If $1 \le p < \infty$ and if f is in $L^p\big([-\pi, \pi], \frac{1}{2\pi}\,d\theta\big)$, prove that the Poisson integral $u(r, \theta)$ of f has the properties that $\|u(r,\,\cdot\,)\|_p \le \|f\|_p$ for $0 \le r < 1$ and that $u(r,\,\cdot\,)$ tends to f in L^p in the sense that $\lim_{r\uparrow 1} \|u(r,\,\cdot\,) - f\|_p = 0$.

8. Suppose that $1 < p' \le \infty$ and that $u(r, \theta)$ is a harmonic function on the open unit disk such that $\sup_{0 \le r < 1} \|u(r,\,\cdot\,)\|_{p'}$ is finite. By using Problem 13 at the end of Chapter IV and taking a weak-star limit of a suitable sequence of functions $u(r_n, \theta)$ with $\{r_n\}$ increasing to 1, prove that $u(r, \theta)$ is the Poisson integral of a function in $L^{p'}\big([-\pi, \pi], \frac{1}{2\pi}\,d\theta\big)$.

Problems 9–12 concern decomposing any bounded nonnegative additive set function on an algebra into a completely additive part and a "purely finitely additive" part. They make use of Zorn's Lemma (Section A9 of the appendix). A bounded nonnegative additive set function μ will be called **purely finitely additive** if there is no nonzero completely additive set function ν such that $0 \le \nu(E) \le \mu(E)$ for all E.

9. Suppose that μ is an additive set function on the σ-algebra of all subsets of the integers such that μ has image $\{0, 1\}$ and $\mu(\{n\}) = 0$ for every integer n. Prove that μ is purely finitely additive. (Such a μ was constructed by means of a nontrivial ultrafilter in Problems 39–41 at the end of Chapter V.)
10. Use Zorn's Lemma to show that any bounded nonnegative additive set function is the sum of a nonnegative completely additive set function and a purely finitely additive set function.
11. Prove that if ν is a bounded nonnegative completely additive set function and if μ is bounded nonnegative and purely finitely additive with $0 \le \mu(E) \le \nu(E)$ for all E, then $\mu = 0$.
12. Deduce from the previous problem and the Jordan Decomposition Theorem that the decomposition of Problem 10 is unique.

Problems 13–15 prove the theorem, for the case of $\mathbb{R}^2$, of Jessen–Marcinkiewicz–Zygmund concerning strong differentiation of integrals of L^p functions almost everywhere when $p > 1$. Strong differentiation holds at (x, y) for the locally integrable function f on $\mathbb{R}^2$ if

$$\lim_{\substack{\text{diam}(R)\to 0,\\ R=\text{geometric rectangle}\\ \text{centered at } (x,y)}} \frac{1}{m(R)} \int_R f(u, v)\, dv\, du = f(x, y).$$

Let f^{**} be the associated maximal function, given by

$$f^{**}(x, y) = \sup_{\substack{\text{diam}(R)\to 0,\\ R=\text{geometric rectangle}\\ \text{centered at } (x,y)}} \frac{1}{m(R)} \int_R |f(u, v)|\, dv\, du.$$

13. Let $f_1(x, y)$ be the value of the one-dimensional Hardy–Littlewood maximal function of $y \mapsto f(x, y)$, and let $f_2(x, y)$ be the value of the one-dimensional Hardy–Littlewood maximal function of $x \mapsto f_1(x, y)$. Prove that $f^{**}(x, y) \le f_2(x, y)$.
14. Using Corollary 9.21 and the previous problem, prove that $\|f^{**}\|_p \le A_p^2 \|f\|_p$ if $1 < p \le \infty$.
15. Conclude that strong differentiation holds almost everywhere for each f in $L^p(\mathbb{R}^2)$ if $1 < p \le \infty$.

Problems 16–19 concern the Hilbert transform H defined in Section VIII.7 and Theorem 9.23. The operator H is defined on $L^p(\mathbb{R}^1)$ for $1 < p < \infty$. Recall the functions h_ε, Q_ε, and ψ_ε on $\mathbb{R}^1$ satisfying $Q_\varepsilon = h_\varepsilon + \psi_\varepsilon$. Let f be in L^p, and let f^* be the Hardy–Littlewood maximal function of f.

16. Prove that there exists a continuous integrable function $\Phi \geq 0$ on $\mathbb{R}^1$ of the form $\Phi(x) = \Phi_0(|x|)$, where Φ_0 is a decreasing C^1 function on $[0, \infty)$, such that the function ψ_ε for $\varepsilon = 1$ satisfies $|\psi_1| \leq \Phi$.

17. Deduce from the previous problem and Corollary 6.42 that $\sup_{\varepsilon>0} |(\psi_\varepsilon * f)(x)| \leq Cf^*(x)$. How does it follow that $\lim_{\varepsilon\downarrow 0}(\psi_\varepsilon * f)(x) = 0$ almost everywhere for all f in L^p, $1 \leq p \leq \infty$?

18. Prove that $Q_\varepsilon * f = P_\varepsilon * (Hf)$ for $f \in L^p$ with $1 < p < \infty$, where $P_\varepsilon(x) = P(x, \varepsilon)$ is the Poisson kernel.

19. Deduce from the previous two problems that the limit in the equality

$$Hf(x) = \lim_{\varepsilon\downarrow 0} \frac{1}{\pi} \int_{|t|\geq\varepsilon} \frac{f(x-t)\,dt}{t}$$

of Theorem 9.23 may be interpreted as an almost-everywhere limit if f is in $L^p(\mathbb{R}^1)$ and $1 < p < \infty$.

Problems 20–22 prove the theorem of M. Riesz that the partial sums of the Fourier series of a function in $L^p([-\pi, \pi])$ converge to the function in L^p if $1 < p < \infty$. Recall from Sections I.10 and VI.7 that if f is integrable on $[-\pi, \pi]$, then the n^{th} partial sum of the Fourier series of f is given by $(S_n f)(x) = (D_n * f)(x)$, where D_n is the Dirichlet kernel $D_n(t) = \frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}$ and the convolution is taken relative to $\frac{1}{2\pi}\,dt$.

20. Suppose it can be proved that $\|S_n f\|_p \leq A_p \|f\|_p$ for $1 < p < \infty$ with A_p independent of n and f. Prove that $S_n f \to f$ in L^p for all f in L^p, provided $1 < p < \infty$.

21. Define $E_n(t) = \frac{2\sin(n+\frac{1}{2})t}{t}$ for $\frac{1}{2n+1} \leq |t| \leq \pi$ and $E_n(t) = 0$ for $|t| < \frac{1}{2n+1}$. Then extend $E_n(t)$ periodically. Show that $D_n - E_n = \varphi_n$ is integrable on $[-\pi, \pi]$ with $\|\varphi_n\|_1 \leq C$ independently of n, and say why it is therefore enough to prove that the operators T_n with $T_n f = E_n * f$ satisfy $\|T_n f\|_p \leq B_p \|f\|_p$ for $1 < p < \infty$ with B_p independent of n and f.

22. In $E_n(t)$, write $\sin(n + \frac{1}{2})t$ as a linear combination of two exponentials e^{ikt}, rewrite each exponential as $e^{-ik(x-t)}e^{ikx}$, and decompose the operator T_n as the corresponding sum of two operators. By relating these two operators separately to the operators H_ε in Theorem 9.23, prove that the T_n's satisfy the desired estimate $\|T_n f\|_p \leq B_p \|f\|_p$.

Problems 23–26 develop a kind of function-valued integration known as **conditional expectation** in probability theory. They make use of the Radon–Nikodym Theorem (Theorem 9.16). Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) = 1$.

23. If f is integrable and if $\mathcal{B}$ is a σ-algebra contained in $\mathcal{A}$, prove that there exists a function $E[f|\mathcal{B}]$ that

    (i) is measurable with respect to $\mathcal{B}$ and
    (ii) has $\int_B f\,d\mu = \int_B E[f|\mathcal{B}]\,d\mu$ for all B in $\mathcal{B}$.

    Show further that $E[f|\mathcal{B}]$ is unique in this sense: any two functions satisfying (i) and (ii) differ only on a set in $\mathcal{B}$ of μ measure 0.

24. Suppose that X is a countable disjoint union of sets X_n in $\mathcal{A}$ and that $\mathcal{B}$ consists of all possible unions of the X_n's. Give an explicit formula for $E[f|\mathcal{B}]$.

25. Show that if $\mathcal{B} = \mathcal{A}$, then $E[f|\mathcal{B}] = f$ almost everywhere.

26. Let $\mathcal{B}$ and $\mathcal{C}$ be σ-algebras with $\mathcal{C} \subseteq \mathcal{B} \subseteq \mathcal{A}$. Prove the following:
    (a) $E[E[f|\mathcal{B}] \mid \mathcal{C}] = E[f|\mathcal{C}]$ almost everywhere.
    (b) If f and g are integrable and everywhere finite, then
    $$E[f+g \mid \mathcal{B}] = E[f|\mathcal{B}] + E[g|\mathcal{B}]$$
    almost everywhere.
    (c) If g is measurable with respect to $\mathcal{B}$ and if f and fg are integrable, then $E[fg \mid \mathcal{B}] = gE[f|\mathcal{B}]$ almost everywhere.
    (d) If f and g are in $L^2(X, \mathcal{A}, \mu)$, then $\int_X f\, E[g|\mathcal{B}]\,d\mu = \int_X E[f|\mathcal{B}]\,g\,d\mu$.

CHAPTER X

Topological Spaces

Abstract. This chapter extends considerably the framework for discussing convergence, limits, and continuity that was developed in Chapter II: topological spaces replace metric spaces.

Section 1 makes various definitions, including definitions for the terms topology, open set, closed set, continuous function, base for a topology, separable, and subspace. It introduces two general kinds of constructions useful in analysis and other fields for forming new topological spaces out of old ones—weak topologies and quotient topologies. The section gives several examples of each.

Sections 2–3 develop standard facts, mostly elementary, about how certain combinations of properties of topological spaces imply others. Examples show some limitations to such implications. Properties that are studied include Hausdorff, regular, normal, dense, compact, locally compact, Lindelöf, and σ-compact.

Section 4 discusses product topologies on arbitrary product spaces, an example of a weak topology. The main theorem, the Tychonoff Product Theorem, says that the product of compact spaces is compact.

Section 5 introduces nets, a generalization of sequences. Sequences by themselves are inadequate for detecting convergence in general topological spaces, and nets are a substitute. The use of nets in many cases provides an easier way of establishing properties of subsets of a topological space than direct arguments with open and closed sets.

Section 6 elaborates on quotient topologies as introduced in Section 1. Conditions under which a quotient space is Hausdorff are of particular interest.

Sections 7–8 prove and apply Urysohn's Lemma, which says that any two disjoint closed sets in a normal topological space may be separated by a real-valued continuous function. This result is fundamental to serious uses of topological spaces in analysis. One application is to showing that every separable Hausdorff regular topology arises from a metric.

Section 9 extends Ascoli's Theorem and the Stone–Weierstrass Theorem from their settings in compact metric spaces in Chapter II to the wider setting of compact Hausdorff spaces.

1. Open Sets and Constructions of Topologies

In applications involving metric spaces, we have seen several times that the explicit form of a metric may not at all be one of objects of interest for the space. Instead, we may be interested in the open sets, or in convergence, or in continuity, or in some other aspect of the space. The same open sets, convergence, and continuity may come from two different metrics, and we have even encountered notions of convergence that are not associated with any metric at all. We saw in Section II.5, for example, that we could associate three different natural-looking metrics to the product $X \times Y$ of two metric spaces, and the three metrics led to the same open sets, the same convergence of sequences, and the same continuous functions. On the other hand, the notions in Chapter V of pointwise convergence, convergence almost everywhere, and weak-star convergence were defined without reference to a metric, and depending on the details of the situation, there need not be metrics yielding these notions of convergence. We have brushed against further, more subtle situations with one or the other of these phenomena—no special distinguished metric or no metric at all—but there is no need to produce a complete list. The present chapter introduces and studies an abstract generalization of the notion of a metric space, namely a "topology," that makes it unnecessary to have the kind of explicit formula demanded by the definition of metric space.

The framework for a "topological space" consists of a nonempty set and a collection of "open sets" satisfying the conditions of Proposition 2.5. Thus let X be a nonempty set. A set $\mathcal{T}$ of subsets of X is called a **topology** for X if

(i) X and $\varnothing$ are in $\mathcal{T}$,
(ii) any union of members of $\mathcal{T}$ is a member of $\mathcal{T}$,
(iii) any finite intersection of members of $\mathcal{T}$ is a member of $\mathcal{T}$.

The members of $\mathcal{T}$ are called **open sets**, and $(X, \mathcal{T})$ is called a **topological space**. When there is no chance for ambiguity, we may refer to X itself as a topological space.

Every metric space furnishes an example of a topological space by virtue of Proposition 2.5; we refer to the topology in question as the **metric topology** for the space. Two other examples of general constructions leading to topological spaces will be given later in this section, and some specific examples of other kinds will be given in Section 2.

Neighborhoods, open neighborhoods, interior, closed sets, limit points, and closure may be defined in the same way as in Section II.2. As remarked after Corollary 2.11, the proofs of certain results relating these notions depended only on the definitions and the three properties of open sets listed above. These results are Proposition 2.6 and Corollary 2.7 characterizing interior, Proposition 2.8 giving properties of the family of all closed sets, Proposition 2.9 relating closed sets to limit points, and Proposition 2.10 and Corollary 2.11 characterizing closure. Thus we may take all those results as known for general topological spaces, and it is not necessary to repeat their statements here.

The notion of continuity extends to topological spaces in straightforward fashion. Specifically the definition of continuity at a point is extracted from the statement of Proposition 2.13: if X and Y are topological spaces, a function

$X \to Y$ is **continuous at a point** $x \in X$ if for any open neighborhood V of $f(x)$ in Y, there is a neighborhood U of x such that $f(U) \subseteq V$. Then Corollary 2.14 is immediately available, saying that if $f : X \to Y$ is continuous at x and $g : Y \to Z$ is continuous at $f(x)$, then the composition $g \circ f$ is continuous at x.

Proposition 2.15 and its proof are available also, saying that the function $f : X \to Y$ is continuous at every point of X if and only if the inverse image under f of every open set in Y is open in X, if and only if the inverse image under f of every closed set in Y is closed in X. We say that $f : X \to Y$ is **continuous** if these equivalent conditions are satisfied. The function $f : X \to Y$ is said to be a **homeomorphism** if f is continuous, f is one-one and onto, and $f^{-1} : Y \to X$ is continuous. The relation "is homeomorphic to" is an equivalence relation.

Now let us come to the two general constructions of topological spaces, known as "weak topologies" and "quotient topologies." Both of these have many applications in real analysis.

The notion of "weak topology" starts from the fact that the intersection of a nonempty collection of topologies for a set is a topology; this fact is evident from the very definition. The prototype of a weak topology is the "product topology" for the product of a nonempty set of topological spaces. In the terminology of Section A1 of the appendix, if S is a nonempty set and if X_s is a nonempty set for each s in S, then the Cartesian product $X = \bigtimes_{s \in S} X_s$ is the set of all functions f from S into $\bigcup_{s \in S} X_s$ such that $f(s)$ is in X_s for all $s \in S$. Now suppose that each X_s is a topological space, and let $p_s : X \to X_s$ be the s^{th} coordinate function, given by $p_s(f) = f(s)$. If X is given the **discrete topology** $\mathcal{D}$, in which every subset of X is open, then each p_s is continuous; in fact, the inverse image of an open set in X_s is some subset of X, and every subset of X is in $\mathcal{D}$. Form the collection of all topologies $\mathcal{T}_\alpha$ on X such that each $p_s : X \to X_s$ is continuous relative to $\mathcal{T}_\alpha$. The collection is nonempty since $\mathcal{D}$ is one. Let $\mathcal{T}$ be their intersection. The inverse image of any open set in X_s under p_s lies in $\mathcal{T}_\alpha$ for each α and hence lies in $\mathcal{T}$. Therefore each p_s is continuous relative to $\mathcal{T}$. We speak of $\mathcal{T}$ as the "weakest topology" on X such that all p_s are continuous, and this topology for X is called the **product topology** for X. We shall study product topologies in more detail in Section 4.

More generally let X be a nonempty set, let S be a nonempty set, let X_s be a topological space for each s in S, and suppose that we are given a function $f_s : X \to X_s$ for each s in S. If X is given the discrete topology, then every f_s is continuous. Arguing as in the previous paragraph, we see that there exists a smallest topology for X making all the functions f_s continuous. This is called the **weak topology for X determined by** $\{f_s\}_{s \in S}$.

EXAMPLES.

(1) Let (X, d) be a metric space. Then the weak topology for X determined by all functions $x \mapsto d(x, y)$ as y varies through X is the usual metric topology on X, as we readily check from the definitions.

(2) Let X be a normed linear space with field of scalars $\mathbb{F}$, such as an L^p space for $1 \le p \le \infty$, and let X^* be the vector space of continuous linear functionals on X, as introduced in Section V.9. (For $X = L^p$ with $1 \le p < \infty$ and with the assumption that the underlying measure is σ-finite, Theorem 9.19 identified X^* explicitly as $L^{p'}$, where p' is the dual index to p.) Each member x of X defines a function $f_x : X^* \to \mathbb{F}$ by the formula $f_x(x^*) = x^*(x)$. The weak topology on X^* determined by X is called the **weak-star topology** on X^* relative to X. The words "relative to X" are included in the terminology because two normed linear spaces X might have the same set X^* of continuous linear functionals. In Section V.9 we introduced a notion of weak-star convergence but no metric associated to it. In problems at the ends of Chapters VI, VIII, and IX, this kind of convergence became a powerful tool for working with harmonic functions, Poisson integrals, and positive definite functions. Later in the present chapter we shall relate topologies to convergence of sequences,[1] and it will be apparent that weak-star convergence as defined in Section V.9 is the appropriate notion of convergence for the newly defined weak-star topology.

(3) The construction in Example 2 can be transposed to other situations in which a topology is to be imposed on a vector space. For example, let X be a normed linear space with field of scalars $\mathbb{F}$ equal to $\mathbb{R}$ or $\mathbb{C}$, and let X^* be the vector space of continuous linear functionals on X. Then X^* indexes a set of functions $x^* : X \to \mathbb{F}$. The weak topology on X determined by X^* is known as the **weak topology** on X. This topology arises in some advanced situations, but we shall not have occasion to make use of it in the present volume.

(4) We have encountered three vector spaces of scalar-valued smooth functions on open sets of Euclidean space—in Section III.2 the space $C^\infty(U)$ of all smooth functions on U, in Section VIII.4 the space $C^\infty_{\text{com}}(U)$ of all smooth functions on U with compact support contained in U, and in Section VIII.4 the space $\mathcal{S}(\mathbb{R}^N)$ of Schwartz functions defined on $\mathbb{R}^N$. The subject of partial differential equations makes extensive use of functions of all three of these kinds, and it is necessary to be able to discuss convergence for them. The easiest convergence to describe is for $C^\infty(U)$, where convergence is to mean uniform convergence of the function and all of its partial derivatives on each compact subset of U. Uniform convergence by itself is captured by the supremum norm, and somehow we want to work here with the supremum norms of the function and each of its partial derivatives on each compact subset. The appropriate topology turns out to be the weak topology determined by all the functions $f \mapsto \|f - g\|$, where $\|\cdot\|$ is the supremum of some iterated partial derivative on some compact subset of U. This construction is carried out in detail in the companion volume *Advanced Real Analysis*. A topology for the Schwartz space $\mathcal{S}(\mathbb{R}^N)$ is obtained in a qualitatively similar way.

[1] And to "nets," which are a generalization of sequences.

A topology for $C^\infty_{\text{com}}(U)$ is more subtle, and it too is constructed in the companion volume.

The second general construction of topological spaces is the "quotient topology" for the set of equivalence classes on X when X is a topological space and some equivalence relation[2] has been specified on X. If the relation is written as $\sim$, the set of equivalence classes may be written as $X/\sim$, and the **quotient map**, i.e., passage from each member of X to its equivalence class, is a well-defined function $q : X \to X/\sim$. With a topology in place on X, define a subset U of $X/\sim$ to be open if $q^{-1}(U)$ is open. Since inverse images of functions preserve set-theoretic operations, it is immediate that the resulting collection of open subsets of $X/\sim$ is a topology for $X/\sim$ and that this topology makes q continuous. This topology is called the **quotient topology** for $X/\sim$. In any other topology $\mathcal{T}'$ on $X/\sim$, any subset V of $X/\sim$ that is open in $\mathcal{T}'$ but not open in the quotient topology must have the property that $q^{-1}(V)$ is not open; this condition implies that q is not continuous when $\mathcal{T}'$ is the topology on $X/\sim$. Therefore the quotient topology is the **finest topology** on $X/\sim$ that makes the quotient map continuous—in the sense that it contains all topologies making q continuous.

EXAMPLES.

(1) Let (X, d) be a pseudometric space such as the set of all integrable functions on some measure space $(S, \mathcal{A}, \mu)$ with $d(g, h) = \int_S |g - h|\, d\mu$. The pseudometric on X gives X a topology. For x and y in X, define $x \sim y$ if $d(x, y) = 0$. The result is an equivalence relation, and we know from Proposition 2.12 that the pseudometric d descends to be a metric on the set $X/\sim$ of equivalence classes. The quotient topology on $X/\sim$ coincides with the topology defined by this metric.

(2) Let X be the interval $[-\pi, \pi]$ with its usual topology from the metric on $\mathbb{R}$, let S^1 be the unit circle in $\mathbb{C}$ with its usual topology from the metric on $\mathbb{C}$, and let $q : X \to S^1$ be given by $q(x) = e^{ix}$. We can consider S^1 as the set of equivalence classes of X under the relation that lets $-\pi$ and π be the only nontrivial pair of elements of X that are equivalent. The function q is continuous, and it carries compact sets to compact sets. In Problem 11 at the end of the chapter, we shall see that q exhibits S^1 as having the quotient topology.

(3) Let X be the line $\mathbb{R}$ with its usual metric, let S^1 be the unit circle as in the previous example, and let $q : X \to S^1$ be given by $q(x) = e^{ix}$. The domain X is a group, and the function q identifies S^1 set-theoretically as the quotient group $\mathbb{R}/2\pi\mathbb{Z}$, where $\mathbb{Z}$ is the subgroup of integers. This example illustrates the natural

[2]Equivalence relations and their connection with equivalence classes are discussed in Section A6 of the appendix.

topology to impose on any quotient of a group when the group has a topology for which all translations are homeomorphisms.[3]

In many situations the problem of describing what sets are to be open sets in a topological space is simplified by the notion of a base for a topology. By a **base** $\mathcal{B}$ for the topology $\mathcal{T}$ on X is meant a subfamily of members of $\mathcal{T}$ such that every member of $\mathcal{T}$ is a union of sets in $\mathcal{B}$. In Chapter II the topology for a metric space was really introduced by specifying that the family of all open balls is to be a base. Arguing as with Proposition 2.31, we obtain the following result.

Proposition 10.1. A family $\mathcal{B}$ of subsets of a nonempty set X is a base for some topology $\mathcal{T}$ on X if and only if

(a) $X = \bigcup_{B\in\mathcal{B}} B$ and
(b) whenever U and V are in $\mathcal{B}$ and x is in $U \cap V$, then there is a B in $\mathcal{B}$ such that x is in B and $B \subseteq U \cap V$.

In this case the topology $\mathcal{T}$ is necessarily the set of all unions of members of $\mathcal{B}$, and hence $\mathcal{T}$ is determined by $\mathcal{B}$. A family $\mathcal{B}$ of subsets of X is a base for a particular given topology $\mathcal{T}_0$ on X if and only if (a) holds and

(b′) for each $x \in X$ and member U of $\mathcal{T}_0$ containing x, there is some member B of $\mathcal{B}$ such that x is in B and B is contained in U.

REMARK. Condition (b) is satisfied if $\mathcal{B}$ is closed under finite intersections. Thus any family of subsets of X that is closed under finite intersections and has union X is a base for some topology on X.

A topological space $(X, \mathcal{T})$ is said to be **separable** if $\mathcal{T}$ has a base consisting of only countably many sets.[4] A separable metric space has a countable base consisting entirely of open balls.

As with metric spaces, there is a natural definition of subspaces for general topological spaces. If $(X, \mathcal{T})$ is a topological space and if A is a nonempty subset of X, then the **relative topology** for A is the family of all sets $U \cap A$ with U in $\mathcal{T}$. We can write $\mathcal{T} \cap A$ for this family. It is a simple matter to check that $\mathcal{T} \cap A$ is indeed a topology for A, and we say that $(A, \mathcal{T} \cap A)$ is a **topological subspace** of $(X, \mathcal{T})$. If there is no possibility of confusion and if the relative topology is understood, we may say that "A is a subspace of X."

[3]The definition of "topological group," which is given in the companion volume *Advanced Real Analysis*, imposes further conditions beyond the fact that every translation is a homeomorphism.

[4]Some authors use the word "separable" to mean that X has a countable dense set, but the meaning in the text here is becoming more and more common. The existence of a countable dense set is not a particularly useful property for a general topological space.

Proposition 10.2. If A and B are subspaces of a topological space X with $B \subseteq A \subseteq X$, then the relative topology of B considered as a subspace of X is identical to the relative topology of B considered as a subspace of A.

PROOF. The relative topology of B considered as a subspace of X consists of all sets $U \cap B$ with U open in X, and the relative topology of B considered as a subspace of A consists of all sets $(U \cap A) \cap B$ with U open in X. Thus the result follows from the identity $(U \cap A) \cap B = U \cap (A \cap B) = U \cap B$.

The next two propositions are proved in the same way as Proposition 2.26 and Corollary 2.27.

Proposition 10.3. If A is a subspace of a topological space X, then the closed sets of A are all sets $F \cap A$, where F is closed in X. Consequently B is closed in A if and only if $B = B^{\text{cl}} \cap A$.

Proposition 10.4. If X and Y are topological spaces and $f : X \to Y$ is continuous at a point a of a subspace A of X, then the restriction $f\big|_A : A \to Y$ is continuous at a. Also, f is continuous at a if and only if the function $f_0 : X \to f(X)$ obtained by redefining the range to be the image is continuous at a.

2. Properties of Topological Spaces

Proposition 2.30 listed certain properties of metric spaces as "separation properties." These properties are not shared by all topological spaces, and instead we list them in this section as definitions. After giving the definitions, we shall examine implications among them and some roles that they play. The disproofs of certain implications provide an opportunity to introduce some further examples of topological spaces beyond those obtained from the constructions in Section 1.

Let $(X, \mathcal{T})$ be a topological space. We say that

(i) X is a $\mathbf{T}_1$ **space** if every one-point set in X is closed,
(ii) X is **Hausdorff** if for any two distinct points x and y of X, there are disjoint open sets U and V with $x \in U$ and $y \in V$,
(iii) X is **regular** if for any point $x \in X$ and any closed set $F \subseteq X$ with $x \notin F$, there are disjoint open sets U and V with $x \in U$ and $F \subseteq V$,
(iv) X is **normal** if for any two disjoint closed subsets E and F of X, there are disjoint open sets U and V such that $E \subseteq U$ and $F \subseteq V$.

Proposition 2.30 listed one further property of an arbitrary metric space X, namely that any two disjoint closed sets can be separated by a continuous function from X into $[0, 1]$. Urysohn's Lemma in Section 7 will establish this property for any normal topological space.

Proposition 10.5. If $(X, \mathcal{T})$ is a topological space, then

(a) X is $\mathbf{T}_1$ if and only if for any pair of distinct points x and y, there are open sets U and V such that $x \in U$, $y \notin U$, $x \notin V$, and $y \in V$,

(b) X is regular if and only if for any point x and any closed set F with $x \notin F$, there is an open set U such that $x \in U$ and $U^{\mathrm{cl}} \cap F = \varnothing$,

(c) X is normal if and only if for any pair of disjoint closed sets E and F, there is an open set U such that $E \subseteq U$ and $U^{\mathrm{cl}} \cap F = \varnothing$.

PROOF. If X is $\mathbf{T}_1$ and if x and y are given, we can choose $U = \{y\}^c$ and $V = \{x\}^c$. In the reverse direction, if x is given, choose, for each $y \neq x$, an open set V_y such that $x \notin V_y$ and $y \in V_y$; then $\{x\}^c = \bigcup_y V_y$ is open, and hence $\{x\}$ is closed.

If X is regular and if x and F are given, we can choose disjoint open sets U and V with $x \in U$ and $F \subseteq V$. Then the closed set V^c has $V^c \supseteq U$ and $V^c \cap F = \varnothing$; therefore also $V^c \supseteq U^{\mathrm{cl}}$ and $U^{\mathrm{cl}} \cap F = \varnothing$. In the reverse direction, suppose that x and F are given and that U is an open set with $x \in U$ and $U^{\mathrm{cl}} \cap F = \varnothing$; choosing $V = (U^{\mathrm{cl}})^c$, we see that $x \in U$, $F \subseteq V$, and $U \cap V = \varnothing$.

If X is normal and if E and F are given, we can choose disjoint open sets U and V with $E \subseteq U$ and $F \subseteq V$. Then the closed set V^c has $V^c \supseteq U$ and $V^c \cap F = \varnothing$; therefore also $V^c \supseteq U^{\mathrm{cl}}$ and $U^{\mathrm{cl}} \cap F = \varnothing$. In the reverse direction, suppose that E and F are given and that U is an open set with $E \subseteq U$ and $U^{\mathrm{cl}} \cap F = \varnothing$; choosing $V = (U^{\mathrm{cl}})^c$, we see that $E \subseteq U$, $F \subseteq V$, and $U \cap V = \varnothing$.

Proposition 10.6. If $(X, \mathcal{T})$ is a topological space and

(a) if X is $\mathbf{T}_1$ and normal, then X is regular,

(b) if X is $\mathbf{T}_1$ and regular, then X is Hausdorff,

(c) if X is Hausdorff, then X is $\mathbf{T}_1$.

PROOF. In (a), if x and a disjoint closed set F are given, then $\{x\}$ is closed, and the fact that X is normal implies that we can separate the closed sets $\{x\}$ and F by disjoint open sets. In (b), if x and y are distinct points in X, then $\{y\}$ is closed and the fact that X is regular implies that we can separate the point x and the disjoint closed set $\{y\}$ by disjoint open sets. In (c), the fact that X is Hausdorff means that for any two distinct points x and y, there are disjoint open sets U and V with $x \in U$ and $y \in V$. Then X satisfies the condition in Proposition 10.5a that was shown to be equivalent to the $\mathbf{T}_1$ property.

EXAMPLES.

(1) A space that is not $\mathbf{T}_1$, regular, or normal. Let $X = \{a, b, c\}$, and let $\mathcal{T} = \{\varnothing,\ \{a\},\ \{a, b\},\ \{a, c\},\ \{a, b, c\}\}$.

(2) A space that is $\mathbf{T}_1$ but not Hausdorff. Let X be an infinite set, and let $\mathcal{T}$ consist of the empty set and all complements of finite sets.

(3) A Hausdorff space that is not regular. Let X be the real line. A subset U of X is to be in $\mathcal{T}$ if for each point x of U, there is an open interval I_x containing x such that every rational number in I_x is in U. Then every open interval is in $\mathcal{T}$, and hence X is certainly Hausdorff. On the other hand, the set of rationals is open in this topology, and therefore the set of irrationals is closed. The set of irrationals cannot be separated from the point 0 by disjoint open sets.

(4) A Hausdorff regular space that is not normal. Let X be the closed upper half plane $\{\operatorname{Im} z \geq 0\}$ in $\mathbb{C}$. A base for $\mathcal{T}$ consists of all open disks in X that do not meet the x axis, together with all open disks in X that are tangent to the x axis; the latter sets are to include the point of tangency. It is easy to see that X is Hausdorff, but a little argument is needed to see that X is regular. To begin with, every open set in the usual metric topology for X is in $\mathcal{T}$, and hence every closed set in the usual metric topology for X is closed relative to $\mathcal{T}$. Let p be a point in X, and let F be a $\mathcal{T}$ closed subset of X not containing p. There is no difficulty in separating p and F by disjoint open sets if p has y coordinate positive, and we therefore assume that p lies on the x axis. Since F is closed, Proposition 10.1 produces a basic open set U tangent to the x axis at p such that $U \cap F = \varnothing$. If D denotes a strictly smaller basic open set tangent to the x axis at p, then the only point of the ordinary boundary of U that lies in D^{cl} is p itself. Thus $F \cap D^{\mathrm{cl}} = \varnothing$, and it follows that D and $(D^{\mathrm{cl}})^c$ are disjoint open sets separating p and F. Consequently X is regular. We postpone the argument that X is not normal until Section 7, when Urysohn's Lemma will be available.

(5) A normal space that is not regular. Let $X = \{a, b\}$, and let $\mathcal{T}$ consist of $\varnothing$, $\{a\}$, and $\{a, b\}$.

We shall see in Section 5 that the Hausdorff property is exactly the right condition to make limits be unique, hence to allow a reasonable notion of convergence. Also, in the construction of a quotient space, it is often a subtle matter to decide whether the quotient space is Hausdorff; we shall obtain sufficient conditions in Section 6.

The property of regularity makes possible a generalization of the passage from a pseudometric space of points to a metric space of equivalence classes. The point of departure is the following proposition; we shall examine the resulting quotient space further in Section 6.

Proposition 10.7. Let X be a regular topological space. For points x and y in X, define $x \sim y$ if x is in $\{y\}^c$. Then $\sim$ is an equivalence relation.

PROOF. Certainly x lies in $\{x\}^{\mathrm{cl}}$, and if x lies in $\{y\}^{\mathrm{cl}}$ and y lies in $\{z\}^{\mathrm{cl}}$, then x lies in $\{z\}^{\mathrm{cl}}$. For the symmetry property, we argue by contradiction and use the regularity of X. Suppose that x lies in $\{y\}^{\mathrm{cl}}$ but y does not lie in $\{x\}^{\mathrm{cl}}$. Regularity allows us to find disjoint open sets U and V such that $y \in U$ and $\{x\}^{\mathrm{cl}} \subseteq V$. Then the closed set V^c contains y and hence also $\{y\}^{\mathrm{cl}}$. Since x lies in $\{y\}^{\mathrm{cl}}$, x lies in V^c. But this relationship contradicts the fact that x lies in V. We conclude that $\sim$ is symmetric and is therefore an equivalence relation.

Subspaces of topological spaces inherit certain properties if the original space has them. Among these are $\mathbf{T}_1$, Hausdorff, and separable. A subspace of a normal space need not be normal, as is seen by taking $X = \{a, b, c, d\}$, and $\mathcal{T} = \{\varnothing, \{a\}, \{a, b\}, \{a, c\}, \{a, b, c, d\}\}$, the subspace being $\{a, b, c\}$ and the relatively closed subsets of interest being $\{b\}$ and $\{c\}$. Let us state the result for regularity as a proposition.

Proposition 10.8. A subspace of a regular topological space is regular.

PROOF. Within a subspace A of X, let F be a relatively closed set, and let x be a point of A not in F. By Proposition 10.3 we have $F = F^{\mathrm{cl}} \cap A$, the closure being taken in X. Since x is in A but not F, x is not in F^{cl}. Since X is regular, we can find disjoint open sets U and V in X with $x \in U$ and $F^{\mathrm{cl}} \subseteq V$. Then $U \cap A$ and $V \cap A$ are disjoint relatively open sets containing x and F.

As with metric spaces, a subset D of a topological space X is **dense in** A if $D^{\mathrm{cl}} \supseteq A$; D is **dense** if D is dense in X. A set D is dense if and only if there is some point of D in each nonempty open set of X. If X is separable, then X has a countable dense set; we have only to select one point from each nonempty member of the base.

The properties of bases of a topological space X become more transparent with the aid of the notion of a local base. A set $\mathcal{U}_x$ of open neighborhoods of x is a **local base** at x if each open set containing x contains some member of $\mathcal{U}_x$. If $\mathcal{B}$ is a base, then the members of $\mathcal{B}$ containing x form a local base at x. Conversely if $\mathcal{U}_x$ is a local base for each x, then the union of all the $\mathcal{U}_x$'s is a base. We say that X has a **countable local base** at each point[5] if a countable such $\mathcal{U}_x$ can be chosen for each x in X. Metric spaces have this property; the open balls of rational radii centered at a point form a local base at the point.

[5] Some authors say instead that "X satisfies the first axiom of countability" or "X is first countable" if this condition holds. In the same kind of terminology, one says that "X satisfies the second axiom of countability" or "X is second countable" if X is separable in the sense of Section 1.

EXAMPLE 4, CONTINUED. A space that has a countable dense set and has a countable local base at each point and yet is not separable. As in Example 4 earlier in this section, let X be the closed upper half plane $\{\operatorname{Im} z \geq 0\}$ in $\mathbb{C}$. A base for $\mathcal{T}$ consists of all open disks in X that do not meet the x axis, together with all open disks in X that are tangent to the x axis; the latter sets are to include the point of tangency. For a point p on the x axis, the open disks of rational radii with point of tangency p form a countable local base, and for a point p off the x axis, the open disks within the open upper half plane having center p and rational radius form a countable local base. A countable dense set consists of all points with rational coordinates and with y coordinate positive. We shall see in Corollary 10.10 in the next section that a separable regular space has to be normal, and this X is not normal, according to the statement in Example 4 and the proof to be given in Section 7. Thus X cannot be separable.

3. Compactness and Local Compactness

Let X be a topological space. In this section we carry over to a general topological space X some definitions made in Section II.7 for metric spaces. A collection $\mathcal{U}$ of open sets is an **open cover** of X if its union is X. An **open subcover** of $\mathcal{U}$ is a subset of $\mathcal{U}$ that is itself an open cover.

We begin with a new term, saying that the topological space X is a **Lindelöf space** if every open cover of X has a countable subcover. Proposition 2.32 showed that a metric space X is separable if and only if X is a Lindelöf space. For general topological spaces it is still true that any separable X is a Lindelöf space, by the same argument as for the implication that condition (a) implies condition (b) in Proposition 2.32. In fact, every subspace of a separable space is separable, and hence every subspace of a separable space is Lindelöf. However, a Lindelöf space need not be separable, as the following example shows rather emphatically.

EXAMPLE. We construct a topological space $(X, \mathcal{T})$ that is Hausdorff and normal, has a countable dense set, has a countable local base at each point, is Lindelöf, yet is not separable. Take X as a set to be the real line. The intersection of any two bounded intervals of the form $[a, b)$ is an interval of the same kind, and the union of all such intervals is the whole line. Hence the bounded intervals $[a, b)$ form a base for some topology on the line, and this topology we take to be $\mathcal{T}$. It is called the **half-open interval topology** for the real line. Since every ordinary open interval of the line is the union of intervals $[a, b)$, any open set in the usual metric topology is open in the half-open interval topology. Any two distinct points of X may be separated by ordinary disjoint open intervals, and therefore X is Hausdorff. To see that X is regular, let a point x and a closed set

F with x not in F be given. Since x is in the open set F^c, some $[x, x+\epsilon)$ is disjoint from F. Then $U = [x, x+\epsilon)$ and $V = (-\infty, x) \cup [x+\epsilon, +\infty)$ are disjoint open sets separating x and F, and we conclude that X is regular. Once we prove that X is Lindelöf, it will follow from Proposition 10.9 below that X is normal. The rationals form a countable dense subset of X, and the set of all intervals $\big[x, x+\frac{1}{n}\big)$ is a countable local base at x. The space X is not separable. In fact, if $\mathcal{B}$ is any base, we can find, for each x, some open neighborhood B_x of x that is in $\mathcal{B}$ and is contained in $[x, x+1)$. If $x < y$, then x cannot lie in B_y and hence $B_x \neq B_y$; therefore $\mathcal{B}$ has to be uncountable. Finally let us see that X is Lindelöf. Let an open cover $\mathcal{U}$ of X be given, and fix a negative real number x_0. Consider the set $S(x_0)$ of all real numbers x such that some countable collection of members of $\mathcal{U}$ covers $[x_0, x]$. Since x_0 is covered by some member of $\mathcal{U}$, the set $S(x_0)$ contains x_0. If the set contains an element x_1, then the member of the countable collection that covers x_1 must contain $[x_1, x_1+\epsilon)$ for some $\epsilon > 0$. Thus $x_1 + \frac{\epsilon}{2}$ is in $S(x_0)$, and $S(x_0)$ contains no largest element. We shall show that $S(x_0) = [x_0, +\infty)$. If the contrary is true, then $S(x_0)$ must be bounded. In this case, let c be the least upper bound. For large enough n, $c - \frac{1}{n}$ is in $S(x_0)$. Taking the union of the countable collections that cover $\big[x_0, c-\frac{1}{n}\big]$, together with one more set to cover c, we obtain a countable collection that covers $[x_0, c]$, and we see that c is in $S(x_0)$. Thus c is in $S(x_0)$, and we have a contradiction to the fact that $S(x_0)$ contains no largest element. We conclude that some countable subcollection of $\mathcal{U}$ covers $[x_0, +\infty)$, no matter what x_0 is. Taking the union of the countable subcollections corresponding to each negative integer, we obtain a countable subcollection of $\mathcal{U}$ covering $(-\infty, +\infty)$. Thus X is Lindelöf.

It is not always so obvious when a topological space is normal. The next result provides one sufficient condition.

Proposition 10.9 (Tychonoff's Lemma). Every regular Lindelöf space is normal.

PROOF. Let X be regular and Lindelöf, and let disjoint closed subsets E and F of X be given. By regularity and Proposition 10.5b each point of E has an open neighborhood whose closure is disjoint from F. Therefore the class $\mathcal{U}$ of open sets with closures disjoint from F covers E. Similarly the class $\mathcal{V}$ of open sets with closures disjoint from E covers F. Thus $\mathcal{U} \cup \mathcal{V} \cup \{X - (E \cup F)\}$ is an open cover of X. Since X is Lindelöf, there exist sequences of sets U_n in $\mathcal{U}$ and V_n in $\mathcal{V}$ such that $E \subseteq \bigcup_{n=1}^{\infty} U_n$ and $F \subseteq \bigcup_{n=1}^{\infty} V_n$. Put

$$U_n' = U_n - \bigcup_{k \le n} V_k^{\text{cl}} \qquad \text{and} \qquad V_n' = V_n - \bigcup_{k \le n} U_k^{\text{cl}}.$$

When $m \le n$, we have $V_m \subseteq \bigcup_{k \le n} V_k^{\mathrm{cl}}$. Then $U_n' \cap V_m = \varnothing$, and hence the smaller set $U_n' \cap V_m'$ is empty. Reversing the roles of the U's and the V's shows that $U_n' \cap V_m'$ is empty for $m \ge n$. Therefore $U_n' \cap V_m' = \varnothing$ for all n and m. Define

$$U = \bigcup_{n=1}^{\infty} U_n' \qquad \text{and} \qquad V = \bigcup_{m=1}^{\infty} V_m'.$$

Then $U \cap V = \bigcup_{n,m} (U_n' \cap V_m') = \varnothing$. Also,

$$E \cap U = E \cap \bigcup_{n=1}^{\infty} \Big(U_n - \bigcup_{k \le n} V_k^{\mathrm{cl}} \Big) \supseteq E \cap \bigcup_{n=1}^{\infty} \Big(U_n - \bigcup_{k=1}^{\infty} V_k^{\mathrm{cl}} \Big) = E \cap \Big(X - \bigcup_{k=1}^{\infty} V_k^{\mathrm{cl}} \Big),$$

the last equality holding since $\{U_n\}$ covers E. The right side here equals E since $V_k^{\mathrm{cl}} \subseteq X - E$ for all k, and therefore $E \subseteq U$. Similarly $F \subseteq V$. The proof is complete.

Corollary 10.10. Every regular separable space is normal.

PROOF. A separable space is automatically Lindelöf, and thus the corollary follows from Proposition 10.9.

Let us return to the concluding example in Section 2, in which X as a set is the closed upper half plane $\{\operatorname{Im} z \ge 0\}$ but in which the topology is nonstandard near the real axis. It was shown in Section 2 that this particular X is regular, and it was stated that Urysohn's Lemma would be used in Section 7 to show that X is not normal. By Corollary 10.10, X cannot be separable. This completes the argument that X has a countable dense set and has a countable local base at each point yet is not separable.

We can now proceed with carrying over some definitions from Section II.7, valid there for metric spaces, to a general topological space X. We call X **compact** if every open cover of X has a finite subcover. A subset E of X is **compact** if it is compact as a subspace of X, i.e., if every collection of open sets in X whose union contains E has a finite subcollection whose union contains E. It is immediate from the definition that the union of two compact subsets is compact.

This definition generalizes the property of closed bounded sets of $\mathbb{R}^n$ given by the Heine–Borel Theorem. We shall see that the Heine–Borel property, rather than the Bolzano–Weierstrass property for sequences, is the useful property to carry over to more general situations in real analysis. In fact, in several places in this book, we have combined an iterated application of the Bolzano–Weierstrass property with the Cantor diagonal process to obtain some conclusion. This construction is tantamount to proving that the product of countably many compact metric spaces, which is a metric space essentially by Proposition 10.28 below, is compact. There will be situations for which we want to consider an uncountable product of compact metric spaces, and then arguments with sequences are not decisive. Instead, it is the Heine–Borel property that is relevant. The Tychonoff Product Theorem of Section 4 will be the substitute for the Cantor diagonal process, and the use of nets, considered in Section 5, will be analogous to the use of sequences.

A number of the simpler results in Section II.7 generalize easily from compact metric spaces to all compact topological spaces or at least to all compact Hausdorff spaces. We list those now. A consequence of Proposition 10.12 below is that compactness is preserved under homeomorphisms.

A set of subsets of a nonempty set is said to have the **finite-intersection property** if each intersection of finitely many of the subsets is nonempty.

Proposition 10.11. A topological space X is compact if and only if each set of closed subsets of X with the finite-intersection property has nonempty intersection.

PROOF. Closed sets with the finite-intersection property have complements that are open sets, no finite subcollection of which is an open cover.

Proposition 10.12. Let X and Y be topological spaces with X compact. If $f : X \to Y$ is continuous, then $f(X)$ is a compact subset of Y.

PROOF. If $\{U_\alpha\}$ is an open cover of $f(X)$, then $\{f^{-1}(U_\alpha)\}$ is an open cover of X. Let $\{f^{-1}(U_j)\}_{j=1}^n$ be a finite subcover. Then $\{U_j\}_{j=1}^n$ is a finite subcover of $f(X)$.

Corollary 10.13. Let X be a compact topological space, and let $f : X \to \mathbb{R}$ be a continuous function. Then f attains its maximum and minimum values.

PROOF. By Proposition 10.12, $f(X)$ is a compact subset of $\mathbb{R}$. Arguing as in the proof of Corollary 2.39, we see that $f(X)$ has a finite supremum and a finite infimum and that both of these must lie in $f(X)$.

Proposition 10.14. A closed subset of a compact topological space is compact.

PROOF. Let E be a closed subset of the compact space X, and let $\mathcal{U}$ be an open cover of E. Then $\mathcal{U} \cup \{E^c\}$ is an open cover of X. Passing to a finite subcover and discarding E^c, we obtain a finite subcover of E. Thus E is compact.

Lemma 10.15. Let K and E be subsets of a topological space X, and let K be compact. Suppose that to each point x of K there are disjoint open sets U_x and V_x such that x is in U_x and $E \subseteq V_x$. Then there exist disjoint open sets U and V such that $K \subseteq U$ and $E \subseteq V$.

PROOF. As x varies through K, the open sets U_x form an open cover of K. By compactness, a finite subcollection of the U_x's is a cover, say $U_{x_1}, \dots, U_{x_n}$. Put $U = \bigcup_{k=1}^n U_{x_k}$ and $V = \bigcap_{k=1}^n V_{x_k}$. Then $K \subseteq U$ and $E \subseteq V$. Also, $U \cap V = \big(\bigcup_{k=1}^n U_{x_k}\big) \cap \big(\bigcap_{k=1}^n V_{x_k}\big) = \bigcup_{k=1}^n \big(U_{x_k} \cap \big(\bigcap_{l=1}^n V_{x_l}\big)\big) \subseteq \bigcup_{k=1}^n (U_{x_k} \cap V_{x_k}) = \varnothing$, and thus U and V have the required properties.

Proposition 10.16. Every compact Hausdorff space is regular and normal.

PROOF. Let X be compact Hausdorff. If a point x and a closed set F with $x \notin F$ are given, we observe by Proposition 10.14 that F is compact. The Hausdorff property of X allows us to take $E = \{x\}$ and $K = F$ in Lemma 10.15, and we obtain disjoint open sets U and V such that x is in V and $F \subseteq U$. Thus X is regular.

If disjoint closed sets E and F are given, then F is compact by Proposition 10.14. The fact that X has been shown to be regular allows us to take $K = F$ in Lemma 10.15, and we obtain disjoint open sets U and V such that $E \subseteq V$ and $F \subseteq U$. Thus X is normal.

Proposition 10.17. In a Hausdorff space every compact set is closed.

PROOF. Let X be a Hausdorff space, and let K be a compact subset of X. Fix x in K^c. The Hausdorff property of X allows us to take $E = \{x\}$ in Lemma 10.15, and we obtain disjoint open sets U_x and V_x such that x is in V_x and $K \subseteq U_x$. Letting x now vary, we see that $K^c = \bigcup_{x \in K^c} V_x$. Hence K^c is open and K is closed.

Corollary 10.18. Let X and Y be topological spaces with X compact and with Y Hausdorff. If $f : X \to Y$ is continuous, one-one, and onto, then f is a homeomorphism.

PROOF. We are to show that $f^{-1} : Y \to X$ is continuous. Let E be a closed subset of X, and consider $(f^{-1})^{-1}(E) = f(E)$. The set E is compact in X by Proposition 10.14, $f(E)$ is compact by Proposition 10.12, and $f(E)$ is closed by Proposition 10.17. Since the inverse image under f^{-1} of any closed set is closed, f^{-1} is continuous.

A topological space is **locally compact** if every point has a compact neighborhood. Compact spaces are locally compact, but the real line with its usual topology is locally compact and not compact. In a sense to be made precise in the next two propositions, locally compact Hausdorff spaces are just one point away from being compact Hausdorff.

Let $(X, \mathcal{T})$ be an arbitrary topological space. Define a new set X^* by $X^* = X \cup \{\infty\}$, where ∞ is not already a member of X, and define $\mathcal{T}^*$ to be the union of $\mathcal{T}$ and the set of all complements in X^* of closed compact subsets of X. We shall verify in Proposition 10.19 that $\mathcal{T}^*$ is a topology for X^*. The topological space $(X^*, \mathcal{T}^*)$ is called the **one-point compactification** of $(X, \mathcal{T})$. By way of examples, the one-point compactification of $\mathbb{R}$ may be visualized as a circle and the one-point compactification of $\mathbb{R}^2$ may be visualized as a sphere.

Proposition 10.19. If $(X, \mathcal{T})$ is a topological space, then $(X^*, \mathcal{T}^*)$ is a compact topological space, X is an open subset of X^*, and the relative topology for X in X^* is $\mathcal{T}$.

PROOF. To see that $\mathcal{T}^*$ is a topology, we observe first that $\varnothing$ and X^* are in $\mathcal{T}^*$. If U and V are in $\mathcal{T}^*$, there are three cases in checking that $U \cap V$ is in $\mathcal{T}^*$: If U and V are both in $\mathcal{T}$, then $U \cap V$ is in $\mathcal{T}$ since $\mathcal{T}$ is closed under finite intersections. If U is in $\mathcal{T}$ and V is not, then V^c is closed compact in X, and $X - V^c$ is thus open in X; since $\mathcal{T}$ is closed under finite intersections, $U \cap V = U \cap (X - V^c)$ is in $\mathcal{T}$. If U and V are not in $\mathcal{T}$, then the complements U^c and V^c in X^* are closed compact subsets of X; so is their union $(U \cap V)^c$, and hence $U \cap V$ is in $\mathcal{T}^*$.

We still have to check closure of $\mathcal{T}^*$ under arbitrary unions. Suppose that U_α is in $\mathcal{T}$ for α in an index set A and V_β has closed compact complement for β in an index set B. Then $\bigcup_{\alpha \in A} U_\alpha$ is in $\mathcal{T}$, and if B is nonempty, $\bigcap_{\beta \in B} V_\beta^c$ is a closed subset of one V_β^c and hence is compact; in this case, $\big(\bigcup_{\beta \in B} V_\beta\big)^c$ is closed compact in X, and hence $\bigcup_{\beta \in B} V_\beta$ is in $\mathcal{T}^*$. Thus we have only to check that $U \cup V$ is in $\mathcal{T}^*$ if U is in $\mathcal{T}$ and V^c is closed compact in X. As the intersection of two closed sets, one of which is compact, $(X - U) \cap V^c = (X - U) \cap (X - V)$ is closed and compact in X, and thus $U \cup V = ((X - U) \cap V^c)^c$ is in $\mathcal{T}^*$. Thus $\mathcal{T}^*$ is a topology.

To see that X^* is compact, let $\mathcal{U}$ be an open cover of X^*. Find some V in $\mathcal{U}$ containing the point ∞. The members of $\mathcal{U} \cap \mathcal{T}$ cover the compact subset V^c of X, and there is a finite subcollection $\mathcal{V}$ that covers V^c. Then $\mathcal{V} \cup \{V\}$ is a finite subcollection of $\mathcal{U}$ that covers X^*.

The set X is in $\mathcal{T}$ and is therefore in $\mathcal{T}^*$. Thus X is open in X^*. To complete the proof, we are to show that $\mathcal{T}^* \cap X = \mathcal{T}$. We know that $\mathcal{T}^* \cap X \supseteq \mathcal{T}$. If V is a member of $\mathcal{T}^*$ that does not lie in $\mathcal{T}$, then V^c is closed compact in X, and its complement $X - V^c = V \cap X$ in X is open in X. Hence $V \cap X$ is in $\mathcal{T}$.

Proposition 10.20. If X^* is the one-point compactification of a topological space X, then X^* is Hausdorff if and only if X is both locally compact and Hausdorff.

PROOF. Suppose that X is locally compact and Hausdorff. Since X is Hausdorff, any two points of X can be separated by disjoint open sets in X, and these sets will be open in X^*. To separate a point x in X from ∞, let C be a compact neighborhood of x in X. Since X is Hausdorff, C is closed in X. Thus C^c is in $\mathcal{T}^*$. Then C^o and C^c are disjoint open sets in X^* such that x is in C^o and ∞ is in C^c, and X^* is Hausdorff.

Conversely suppose that X^* is Hausdorff. Proposition 10.19 shows that X is a subspace of X^*. Since any subspace of a Hausdorff space is Hausdorff, X is Hausdorff. To see that X is locally compact, let x be in X, and find disjoint open sets U and V in X^* such that x is in U and ∞ is in V. Then U must be in $\mathcal{T}$, and V^c must be closed compact in X. Since $U \cap V = \varnothing$, $U \subseteq V^c$. This inclusion exhibits V^c as a compact neighborhood of x, and thus X is locally compact.

Corollary 10.21. Every locally compact Hausdorff space is regular.

PROOF. If X is locally compact Hausdorff, Propositions 10.19 and 10.20 show that the one-point compactification X^* is compact Hausdorff and allow us to regard X as a subspace of X^*. Proposition 10.16 shows that X^* is regular, and Proposition 10.8 shows that X is therefore regular.

A locally compact Hausdorff space need not be normal; an example is given in Problem 5 at the end of the chapter. The remainder of this section concerns senses in which a locally compact Hausdorff space is almost normal.

Corollary 10.22. If K and F are disjoint closed sets in a locally compact Hausdorff space and if K is compact, then there exist disjoint open sets U and V such that $K \subseteq U$ and $F \subseteq V$.

PROOF. This is immediate from Lemma 10.15 and Corollary 10.21.

Corollary 10.23. If K is a compact set in a locally compact Hausdorff space, then there is a compact set L such that $K \subseteq L^o$.

PROOF. Let X be locally compact Hausdorff, and form the one-point compactification X^*. Since X^* is compact Hausdorff by Proposition 10.20, Proposition 10.17 shows that K is closed in X^* and Proposition 10.16 shows that X^* is regular. Thus Proposition 10.5b shows that we can find an open set U in X^* such that ∞ is in U and $U^{\text{cl}} \cap K = \varnothing$. Then $K \subseteq X^* - U^{\text{cl}} \subseteq X^* - U$. By definition of the topology of X^*, the set $L = X^* - U$ is compact in X. Its subset $X^* - U^{\text{cl}}$ is open and is therefore contained in L^o. Thus $K \subseteq L^o \subseteq L$ with L compact.

A topological space is called σ-**compact** if there is a sequence of compact sets with union the whole space. The real line with its usual topology is σ-compact. For that matter, so is the subspace of rationals since each finite subset is compact.

Proposition 10.24. A locally compact topological space is σ-compact if and only if it is Lindelöf. Consequently every σ-compact locally compact Hausdorff space is normal.

PROOF. If X is σ-compact, write $X = \bigcup_{n=1}^{\infty} K_n$ with K_n compact. If $\mathcal{U}$ is an open cover of X, then $\mathcal{U}$ is an open cover of each K_n, and there is a finite subcover $\mathcal{U}_n$ of K_n. Then $\bigcup_{n=1}^{\infty} \mathcal{U}_n$ is a countable subcover of $\mathcal{U}$, and X is Lindelöf.

Conversely if X is locally compact and Lindelöf, choose, for each x in X, a compact neighborhood K_x of x, and let U_x be the interior of K_x. As x varies, the U_x form an open cover of X. Since X is Lindelöf, there is a countable subcover $\{U_{x_n}\}_{n=1}^{\infty}$. Since we have $U_{x_n} \subseteq K_{x_n}$ for all n, $\{K_{x_n}\}_{n=1}^{\infty}$ is a sequence of compact sets with union X. Hence X is σ-compact.

Finally if X is locally compact Hausdorff and σ-compact, hence also Lindelöf, then Corollary 10.21 shows that X is regular, and Tychonoff's Lemma (Proposition 10.9) shows that X is normal.

Proposition 10.25. In a σ-compact locally compact Hausdorff space, there exists an increasing sequence $\{K_n\}$ of compact sets with union the whole space and with $K_n \subseteq K_{n+1}^o$ for all n.

PROOF. Let X be a locally compact Hausdorff space such that $X = \bigcup_{n=1}^{\infty} L_n$ with L_n compact. Replacing L_n by the union of the previous members of the sequence, we may assume that $L_n \subseteq L_{n+1}$ for all $n \geq 1$. Put $L_0 = K_0 = \varnothing$. Use Corollary 10.23 to choose K_1 compact with $L_1 \subseteq K_1^0$.

Inductively suppose that $n > 0$ and that for all k with $0 < k \leq n$, a compact set K_k has been defined such that $L_k \cup K_{k-1} \subseteq K_k^o$. Applying Corollary 10.23, we can find a compact set K_{n+1} such that the compact set $L_{n+1} \cup K_n$ is contained in K_{n+1}^o. Then $K_{k-1} \subseteq K_k^o$ for all $k \geq 1$ as required, and $X = \bigcup_{n=1}^{\infty} K_n$ since $K_n \subseteq L_n$ and $\bigcup_{n=1}^{\infty} L_n = X$.

4. Product Spaces and the Tychonoff Product Theorem

The **product topology** for the product of topological spaces was discussed briefly in Section 1. If S is a nonempty set and if X_s is a topological space for each s in S, then the Cartesian product $X = \bigtimes_{s \in S} X_s$, as a set, is the set of all functions f from S into $\bigcup_{s \in S} X_s$ such that $f(s)$ is in X_s for all $s \in S$. The topology that is imposed on X is, by definition, the weakest topology that makes the s^{th} coordinate function $p_s : X \to X_s$ be continuous for every s.

Let us investigate what sets have to be open in this topology, and then we can look at examples and see better what the topology is. If U_s is any open subset of X_s, then $p_s^{-1}(U_s)$ has to be open in X since p_s is continuous. For example, if $S = \{1, 2\}$, we are considering $X = X_1 \times X_2$. A set $p_1^{-1}(U_1)$ is of the form $U_1 \times X_2$, and a set $p_2^{-1}(U_2)$ is of the form $X_1 \times U_2$. These have to be open if U_1 is open in X_1 and U_2 is open in X_2. The intersection of any two such sets, which is of the form $U_1 \times U_2$, has to be open in X, as well. We do not need to intersect these sets further, since $p_1^{-1}(U_1) \cap p_1^{-1}(V_1) = p_1^{-1}(U_1 \cap V_1)$. By the remark with Proposition 10.1, the sets $p_1^{-1}(U_1) \cap p_2^{-1}(U_2)$ with U_1 open in X_1 and U_2 open in X_2 form a base for some topology on $X = X_1 \times X_2$. These sets have to be open in the product topology, and p_1 and p_2 are indeed continuous in this topology. Therefore the product topology on $X = X_1 \times X_2$ has

$$\big\{p_1^{-1}(U_1) \cap p_2^{-1}(U_2) \,\big|\, U_1 \text{ open in } X_1, U_2 \text{ open in } X_2\big\}$$

as a base. More generally the product topology on $X = X_1 \times \cdots \times X_n$ has

$$\Big\{\bigcap_{k=1}^{n} p_k^{-1}(U_k) \,\Big|\, U_k \text{ open in } X_k \text{ for each } k\Big\}$$

as a base.

When the index set S is the set of positive integers, the product $X = \times_{n\in S} X_n$, as a set, is the set of sequences $\{f(n)\}_{n\in S}$. Again any set $p_n^{-1}(U_n)$ with U_n open in X_n has to be open in X. Hence any finite intersection of such sets as n varies has to be open. But there is no need for infinite intersections of such sets to be open, and a base for the product topology in fact consists of all *finite* intersections of sets $p_n^{-1}(U_n)$ with U_n open in X_n.

The use of finite intersections, and not infinite intersections, persists for all S and gives us a description of a base for the product topology in general. When $S = [0, 1]$ and all X_s are $[0, 1]$, the description of the product topology has a helpful geometric interpretation. The set X consists of all functions from the closed unit interval to itself, and we can visualize these in terms of their graphs. A basic open set of such functions imposes restrictions at finitely many values of s, i.e., at finitely many points of the domain. At such values of s, the graph of a function in the basic open set is to pass through a certain window U_s depending on s. At all other values of s, the function is unrestricted.

Proposition 10.26. The topological product of Hausdorff topological spaces is Hausdorff.

PROOF. Let a product $X = \times_{n\in S} X_n$ be given, let $p_s : X \to X_s$ be the s^{th} coordinate function, and let two distinct members f and g of X be given. Members of X are functions of a certain kind, and these two functions, being distinct, have $f(s) \neq g(s)$ for some $s \in S$. Since X_s is Hausdorff, we can choose disjoint open sets U_s and V_s in X_s such that $f(s)$ is in U_s and $g(s)$ is in V_s. Then $p_s^{-1}(U_s)$ and $p_s^{-1}(V_s)$ are disjoint open sets in X such that f is in $p_s^{-1}(U_s)$ and g is in $p_s^{-1}(V_s)$.

Theorem 10.27 (Tychonoff Product Theorem). The topological product of compact topological spaces is compact.

REMARKS. This theorem is a fundamental tool in real analysis. We shall give the proof and then discuss how the theorem can be regarded as a generalization of the Cantor diagonal process used in the proofs earlier of the fact that any totally bounded complete metric space is compact (Theorem 2.46), the Helly Selection Principle (Problem 10 at the end of Chapter I), Ascoli's Theorem (Theorems 1.22 and 2.56), and, by implication, the Cauchy–Peano Existence Theorem for differential equations (Problems 24–29 at the end of Chapter IV). The proof will make use of Zorn's Lemma (Section A9 of the appendix), which is one formulation of the Axiom of Choice. Actually, the Axiom of Choice arises in two more transparent ways in the proof as well. One is simply in the statement that the topological product is a topological space; for this to be the case, the product has to be nonempty, and that is the content of the Axiom of Choice. The other is the construction of a particular element x in the product that occurs near the beginning of the proof below.

PROOF. Let $X = \bigtimes_{s\in S} X_s$ be given with each X_s compact, and let $p_s : X \to X_s$ be the s^{th} coordinate function. We are to prove that any open cover of X has a finite subcover, and we begin by proving a special case. Let $\mathcal{S}$ be the family of all sets $p_s^{-1}(U_s)$ as U_s varies through all open sets of X_s and as s varies. We know that finite intersections of members of $\mathcal{S}$ form a base for the product topology on X. For the special case let $\mathcal{U}$ be an open cover of X by members of $\mathcal{S}$; we shall produce a finite subcover. For each s, let $\mathcal{B}_s$ be the family of all open sets U_s in X_s such that $p_s^{-1}(U_s)$ is in $\mathcal{U}$. We may assume for each s that no finite subfamily of $\mathcal{B}_s$ covers X, since otherwise the corresponding finitely many sets $p_s^{-1}(U_s)$ would cover X. By compactness of X_s, $\mathcal{B}_s$ does not cover X_s; say that x_s is not covered. The point x of X whose s^{th} coordinate is x_s then belongs to no member of $\mathcal{U}$, and $\mathcal{U}$ cannot be a cover. This contradiction shows that the special $\mathcal{U}$ has a finite subcover.

Now let $\mathcal{U}$ be any open cover of X, and suppose that no finite subfamily of $\mathcal{U}$ covers X. Let $\mathcal{C}$ be the system of all open covers $\mathcal{V}$ of X such that $\mathcal{U} \subseteq \mathcal{V}$ and such that no finite subfamily of $\mathcal{V}$ covers X. The set $\mathcal{C}$ is partially ordered by inclusion upward and is nonempty, having $\mathcal{U}$ as a member. If $\{\mathcal{V}_\alpha\}$ is a chain in $\mathcal{C}$, then we shall show that $\mathcal{V} = \bigcup_\alpha \mathcal{V}_\alpha$ is in $\mathcal{C}$ and hence is an upper bound in $\mathcal{C}$ for the chain $\{\mathcal{V}_\alpha\}$. In fact, $\mathcal{V}$ is certainly an open cover. If it has a finite subcover, then each member of the finite subcover lies in one of the covers, say $\mathcal{V}_{\alpha_j}$. Since $\{\mathcal{V}_\alpha\}$ is a chain, all members of the finite subcover lie in the largest of those $\mathcal{V}_{\alpha_j}$'s. Thus one of the $\mathcal{V}_{\alpha_j}$'s fails to be in $\mathcal{C}$, and we arrive at a contradiction. We conclude that every chain in $\mathcal{C}$ has an upper bound in $\mathcal{C}$. By Zorn's Lemma let $\mathcal{U}^*$ be a maximal cover from $\mathcal{C}$ of X.

The family $\mathcal{S} \cap \mathcal{U}^*$ of all members of $\mathcal{U}^*$ that are in the family $\mathcal{S}$ of the first paragraph of the proof has the property that no finite subfamily is a cover of X. By the result of the first paragraph, $\mathcal{S} \cap \mathcal{U}^*$ cannot be a cover of X. Hence we shall have arrived at a contradiction if we show that the union of the members of $\mathcal{U}^*$ is contained in the union of the members of $\mathcal{S} \cap \mathcal{U}^*$. Let U be a member of $\mathcal{U}^*$, and fix a point x in U. Since finite intersections of members of $\mathcal{S}$ form a base, Proposition 10.1 shows that there are members $S_1 \cap \cdots \cap S_n$ of $\mathcal{S}$ such that x is in $S_1 \cap \cdots \cap S_n$ and $S_1 \cap \cdots \cap S_n \subseteq U$. We shall show that one of the sets S_j is in $\mathcal{U}^*$, hence in $\mathcal{U}^* \cap \mathcal{S}$, and then the proof will be complete.

If S_1 is in $\mathcal{U}^*$, we are finished. Otherwise, by the maximality of $\mathcal{U}^*$, there are finitely many open sets $C_1, \ldots, C_k$ of $\mathcal{U}^*$ such that $X = S_1 \cup C_1 \cup \cdots \cup C_k$. Again by the maximality, no open set containing S_1 can belong to $\mathcal{U}^*$, since the union of that set with $C_1 \cup \cdots \cup C_k$ would be X. Proceeding inductively, suppose we have shown that no open set containing $S_1 \cap \cdots \cap S_i$ is in $\mathcal{U}^*$ and that there are open sets $D_1, \ldots, D_m$ in $\mathcal{U}^*$ with

$$X = (S_1 \cap \cdots \cap S_i) \cup (D_1 \cup \cdots \cup D_m).$$

If, as we may assume, S_{i+1} is not in $\mathcal{U}^*$, then by maximality of $\mathcal{U}^*$, there are open sets $E_1, \ldots, E_r$ in $\mathcal{U}^*$ such that $X = S_{i+1} \cup E_1 \cup \cdots \cup E_r$. Then

$$X - S_{i+1} \subseteq E_1 \cup \cdots \cup E_r,$$

and
$$\begin{aligned} S_{i+1} &= (S_1 \cap \cdots \cap S_{i+1}) \cup (S_{i+1} \cap (D_1 \cup \cdots \cup D_m)) \\ &\subseteq (S_1 \cap \cdots \cap S_{i+1}) \cup (D_1 \cup \cdots \cup D_m). \end{aligned}$$

Hence

$$X = S_{i+1} \cup (X - S_{i+1}) \subseteq \big((S_1 \cap \cdots \cap S_{i+1}) \cup (D_1 \cup \cdots \cup D_m)\big) \cup \big(E_1 \cup \cdots \cup E_r\big).$$

That is,

$$X = (S_1 \cap \cdots \cap S_{i+1}) \cup (D_1 \cup \cdots \cup D_m \cup E_1 \cup \cdots \cup E_r).$$

Therefore, once again by maximality of $\mathcal{U}^*$, no open set containing $S_1 \cap \cdots \cap S_{i+1}$ can be in $\mathcal{U}^*$, and the induction is complete. In particular, U, which is an open set containing $S_1 \cap \cdots \cap S_n$, is not in $\mathcal{U}^*$. This contradiction concludes the proof.

As announced above, the Tychonoff Product Theorem is a generalization of the Cantor diagonal process. In fact, let us see how that diagonal process may be used to show directly that the product of a sequence of copies of $[0, 1]$ is compact. Denote the product as a set by $X = \bigtimes_{n=1}^{\infty} [0, 1]$. A member of X is a sequence $\{x_n\}$ with terms x_n. Let us impose on X the Hilbert-cube metric of Example 11 in Section II.1:

$$d(\{x_n\}, \{y_n\}) = \sum_n 2^{-n}|x_n - y_n|.$$

We show below in Corollary 10.29 that this metric on X yields the product topology. By Theorem 2.36 the space X will then be compact if every sequence in X has a convergent subsequence. A sequence in X means a system $\{x_n^{(m)}\}$ in which the n^{th} term of the m^{th} sequence is $x_n^{(m)}$. Convergence is term-by-term convergence. To produce a convergent subsequence of sequences, we iterate use of the Bolzano–Weierstrass property of $[0, 1]$. Remembering that m tells which sequence we are dealing with, we find first a subcollection m_k of the indices m such that we have convergence along the m_k's for $n = 1$, then a subcollection m_{k_l} of that such that we have convergence along the m_{k_l}'s for $n = 2$, and so on. Since the intersection of all these sequences may be empty, we instead obtain a convergent subsequence of our sequences by requiring that the k^{th} term of the desired subsequence be the k^{th} term of the k^{th} subsequence. This "diagonal process" thus shows that any sequence in X has a convergent subsequence. Hence X, being a metric space, is compact.

The general Tychonoff Product Theorem may thus be viewed as a topological generalization of the diagonal process to product spaces with an uncountable number of factors.

Here is one way in which the Tychonoff Product Theorem is used in real analysis. For the situation in which we have a set Y and a system of functions $f_s : Y \to \mathbb{C}$ for s in some set S, the first section of this chapter introduced the **weak topology for Y determined by** $\{f_s\}_{s\in S}$. This is the weakest topology making all the functions f_s continuous. Often in analysis a set Y and a system of functions f_s of this kind arise in a construction, and then this weak topology is imposed on Y. In favorable cases it turns out that each function f_s is bounded on Y. In this case if there are enough functions f_s to **separate points** of Y (i.e., enough so that for each x and y there is some s with $f_s(x) \neq f_s(y)$), then Y is a candidate for a compact Hausdorff space. To see what is needed for compactness, let X_s be a compact subset of $\mathbb{C}$ containing the image of f_s, and let $X = \bigtimes_{s\in S} X_s$. Define a function $F : Y \to X$ by "$F(y)$ is the function whose s^{th} coordinate is $f_s(y)$." It is readily verified that F is a homeomorphism of Y onto a subspace of the compact Hausdorff space X. Thus Y is compact if and only if $F(Y)$ is closed in X. Checking that a set is closed is much easier than checking compactness directly, and it is especially easy if one uses "nets," which are the objects introduced in the next section as a useful generalization of sequences.

To complete our discussion, we still need to prove that the Hilbert-cube metric on $X = \bigtimes_{n=1}^{\infty} [0, 1]$ yields the product topology. It will be helpful to prove the following more general result and to obtain the statement about the Hilbert cube as a special case.

Proposition 10.28. Suppose that X is a nonempty set and $\{d_n\}_{n\geq 1}$ is a sequence of pseudometrics on X such that $d_n(x, y) \leq 1$ for all n and for all x and y in X. Then $d(x, y) = \sum_{n=1}^{\infty} 2^{-n} d_n(x, y)$ is a pseudometric. If the open balls relative to d_n are denoted by $B_n(r; x)$ and the open balls relative to d are denoted by $B(r; x)$, then the B_n's and B's are related as follows:

(a) whenever some $B_n(r_n; x)$ is given with $r_n > 0$, there exists some $B(r; x)$ with $r > 0$ such that $B(r; x) \subseteq B_n(r_n; x)$,

(b) whenever $B(r; x)$ is given with $r > 0$, there exist finitely many $r_n > 0$, say for $n \leq K$, such that $\bigcap_{n=1}^{K} B_n(r_n; x) \subseteq B(r; x)$.

PROOF. For (a), choose $r = 2^{-n} r_n$. If $d(x, y) < r$, then $2^{-m} d_m(x, y) < r$ for all m and in particular $d_n(x, y) < 2^n r = r_n$.

For (b), choose K large enough so that $2^{-K} < r/2$, and put $r_n = r/2$ for $n \leq K$. If y is in $\bigcap_{n=1}^{K} B_n(r_n; x)$, then $d_n(x, y) < r_n = r/2$ for $n \leq K$. Hence $d(x, y) \leq \sum_{n=1}^{K} 2^{-n} d_n(x, y) + \sum_{n=K+1}^{\infty} 2^{-n} < \sum_{n=1}^{K} 2^{-n} r/2 + 2^{-K} < r/2 + r/2 = r$. Therefore y is in $B(r; x)$.

Corollary 10.29. The Hilbert-cube metric on $X = \times_{n=1}^{\infty} [0, 1]$ yields the product topology.

PROOF. Proposition 10.28a implies that any basic open neighborhood of x in the product topology contains a basic open neighborhood in the Hilbert-cube metric topology. Proposition 10.28b shows that any basic open neighborhood of x in the Hilbert-cube metric topology contains a basic open neighborhood in the product topology.

5. Sequences and Nets

Sequences are of limited interest in general topological spaces. Nets, which are generalized sequences of a certain kind, are a useful substitute, and we introduce them in this section. Using nets, we shall be able to see that product topologies are appropriate for detecting pointwise convergence in the same way that the metric topology obtained from the supremum norm is appropriate for detecting uniform convergence.

We begin with two examples that illustrate some of the difficulties with using sequences in general topological spaces. We use the natural definition suggested by Section II.4—that a sequence $\{x_n\}$ in X **converges** to x_0 if for each neighborhood of x_0, there is some N depending on the neighborhood such that x_n is in the neighborhood for $n \geq N$. We say that the sequence is **eventually** in the neighborhood. The point x_0 is a **limit** of the sequence.

EXAMPLES.

(1) Let X be the set of positive integers, and let a topology for X consist of the empty set and all sets whose complements are finite. If $x_n = 2n$, then the sequence $\{x_n\}$ converges to every point of X and hence does not have a unique limit. The space X is $\mathbf{T}_1$ and has a countable local base at each point, but X is not Hausdorff.

(2) Let X be the set of points (m, n) in the plane with m and n integers ≥ 0. Define a topology for X as follows. Any set not containing $(0, 0)$ is to be open. If a set U contains $(0, 0)$, then U is defined to be open if there are only finitely many columns $C_m = \{(m, n) \mid n = 0, 1, 2 \dots\}$ such that $C_m - (U \cap C_m)$ is infinite. Enumerate X, and define x_n to be the n^{th} point in the enumeration. It is easy to check that the image of the sequence $\{x_n\}$ has $(0, 0)$ as a limit point and that no subsequence of $\{x_n\}$ converges to $(0, 0)$. The space X is Hausdorff but does not have a countable local base at $(0, 0)$.

Thus the elementary results in Section II.4 do not generalize to all topological spaces. But Proposition 2.20 (the uniqueness of the limit of any sequence) is still valid if X is Hausdorff, and Proposition 2.22 and Corollary 2.23 (the characterization of limit points and of closed sets in terms of sequences) are still valid if X has a countable local base at each point. Nets will cure the problem about characterizing limit points and closed sets without countable local bases but not the problem about nonuniqueness of limits, and thus we shall be able to work well with nets in all Hausdorff spaces. In particular we shall be able to use nets in uncountable products of Hausdorff spaces, which arise frequently in real analysis and tend not to have a countable local base at each point.

Before defining nets, let us give one positive result whose statement mixes topological spaces and metric spaces. If S is any nonempty set, we have made $B(S)$, the vector space of all bounded scalar-valued functions on S, into a normed linear space—and hence a metric space—by means of the supremum norm. If S is a topological space, let $C(S)$ be the subset of continuous members of $B(S)$; this is a vector subspace and hence is itself a normed linear space.

Proposition 10.30. If S is a topological space and $\{f_n\}$ is a sequence of scalar-valued functions continuous at s_0 and converging uniformly to a function f, then f is continuous at x_0. Consequently the subspace $C(S)$ of $B(S)$ is a closed subspace, and $C(S)$ is complete as a metric space.

PROOF. Given $\epsilon > 0$, choose N such that $n \geq N$ implies $\|f_n - f\|_{\sup} < \epsilon$. For any s, we then have

$$\begin{aligned}|f(s) - f(s_0)| &\le |f(s) - f_N(s)| + |f_N(s) - f_N(s_0)| + |f_N(s_0) - f(s_0)| \\ &\le \|f_N - f\|_{\sup} + |f_N(s) - f_N(s_0)| + \|f_N - f\|_{\sup} \\ &< 2\epsilon + |f_N(s) - f_N(s_0)|.\end{aligned}$$

Since f_N is continuous at s_0, there exists a neighborhood of s_0 such that the right side is $< 3\epsilon$ for s in that neighborhood. Thus f is continuous at s_0.

If $\{f_n\}$ is a sequence in $C(S)$ converging uniformly to f in $B(S)$, then f is in $C(S)$, by the result of the previous paragraph. Since convergence of sequences in $B(S)$ is the same as uniform convergence, Corollary 2.23 shows that $C(S)$ is a closed subset of $B(S)$. Propositions 2.43 and 2.44 then show that $C(S)$ is complete as a metric space.

Now we turn our attention to nets. In the indexing for a net, the set of positive integers is replaced by a "directed set," which we define first. Let D be a partially ordered set in the sense of Section A9 of the appendix, the partial ordering being denoted by $\le$. We say that $(D, \le)$ is a **directed set** if for any α and β in D, there is some γ in D with $\alpha \le \gamma$ and $\beta \le \gamma$.

EXAMPLES.

(1) Take D to be the set of positive integers, and let $\le$ have the usual meaning.

(2) Let S be a nonempty set, take D to be the set of all finite subsets of S, and let $\alpha \le \beta$ mean that the inclusion $\alpha \subseteq \beta$ holds.

(3) Let X be a topological space, let x be a point in X, take D to be the set of all neighborhoods of x, and let $\alpha \le \beta$ mean that $\alpha \supseteq \beta$.

(4) Let $(D_1, \le_1)$ and $(D_2, \le_2)$ be two directed sets, take D to be $D_1 \times D_2$, and let $(\alpha_1, \alpha_2) \le (\beta_1, \beta_2)$ mean that $\alpha_1 \le_1 \beta_1$ and $\alpha_2 \le_2 \beta_2$.

If X is a nonempty set, a **net** in X is a function from a directed set D into X. If D needs to be specified to avoid confusion, we speak of a "net from D to X." The function will often be written $\alpha \mapsto x_\alpha$ or $\{x_\alpha\}$. If E is a subset of X, the net is **eventually in** E if there is some α_0 in D such that $\alpha_0 \le \alpha$ implies that x_α is in E. The net is **frequently in** E if for any α in D, there is a β in D with $\alpha \le \beta$ such that x_β is in E. It is important to observe that the negation of "the net is eventually in E" is that "the net is frequently in the complement of E."

The directedness of the set D plays an important role in the theory by allowing us to work simultaneously with finitely many conditions on a net. For example, if $\{x_\alpha\}$ is eventually in E_1 and eventually in E_2, then it is eventually in $E_1 \cap E_2$. In fact, the given conditions say that there are members α_1 and α_2 of D such that x_α is in E_1 for $\alpha_1 \le \alpha$ and x_α in E_2 for $\alpha_2 \le \alpha$. The directedness implies that

$\alpha_1 \le \alpha_0$ and $\alpha_2 \le \alpha_0$ for some α_0 in D. Then $\{x_\alpha\}$ is in $E_1 \cap E_2$ for $\alpha_0 \le \alpha$. This kind of argument will be used often without mention of the details.

If X is a topological space, a net $\{x_\alpha\}$ in X **converges** to x_0 in X if $\{x_\alpha\}$ is eventually in each neighborhood of x_0. In this case we write $x_\alpha \to x_0$, and we say that x_0 is a **limit** of $\{x_\alpha\}$. Because of the availability of Examples 3 and 4 above, it is an easy matter to characterize the terms "Hausdorff," "limit point," "closed set," and "continuous at a point" in terms of convergence of nets.

Proposition 10.31. A topological space X is Hausdorff if and only if every convergent net in X has only one limit.

PROOF. Suppose that X is Hausdorff and that $x_\alpha \to x_0$ and $x_\alpha \to y_0$ with $x_0 \ne y_0$. Choose disjoint open sets U and V with x_0 in U and y_0 in V. By the assumed convergence, $\{x_\alpha\}$ is in U eventually and is in V eventually. Then it is in $U \cap V = \varnothing$ eventually, and we have a contradiction.

Suppose that X is not Hausdorff. Find distinct points x_0 and y_0 such that every pair of neighborhoods U of x_0 and V of y_0 has nonempty intersection. For any such pair (U, V), define $x_{U,V}$ to be some point in the intersection. Combining Examples 3 and 4 above, we see that $(U, V) \mapsto x_{U,V}$ is a net in X converging to both x_0 and y_0.

Proposition 10.32. If X is a topological space, then

(a) for any subset A of X and limit point x_0 of A, there exists a net in $A-\{x_0\}$ converging to x_0,

(b) any convergent net $\{x_\alpha\}$ in X with limit x_0 in X either has x_0 as a limit point of the image of the net or else is eventually constantly equal to x_0.

PROOF. For (a), the definition of limit point implies that for each neighborhood U of x_0, the set $U \cap (A-\{x_0\})$ is nonempty. If x_U denotes a point in the intersection, then $U \mapsto x_U$ is a net in $A - \{x_0\}$ converging to x_0.

For (b), suppose that x_0 is not a limit point of the image of the net. Then there exists a neighborhood U of x_0 such that $U - \{x_0\}$ is disjoint from the image of the net. Since the convergence implies that the net is eventually in U, it must be true that $x_\alpha = x_0$ eventually.

Corollary 10.33. If X is a topological space, then a subset F of X is closed if and only if every convergent net in F has its limit in F.

PROOF. Suppose that F is closed and that $\{x_\alpha\}$ is a convergent net in F with limit x_0. By Proposition 10.32b, either x_0 is in the image of the net or x_0 is a limit point of the image of the net. In the latter case, x_0 is a limit point of the larger set F. In either case, x_0 is in F; thus the limit of any convergent net in F is in F.

Conversely suppose every convergent net in F has its limit in F. If x_0 is a limit point of F, then Proposition 10.32a produces a net in $F - \{x_0\}$ converging to x_0. By assumption, the limit x_0 is in F. Therefore F contains all its limit points and is closed.

Proposition 10.34. Let $f : X \to Y$ be a function between topological spaces. Then f is continuous at a point x_0 in X if and only if whenever $\{x_\alpha\}$ is a convergent net in X with limit x_0, then $\{f(x_\alpha)\}$ is convergent in Y with limit $f(x_0)$.

REMARKS. This result needs to be used with caution if Y is not known to be Hausdorff. For example, let X and Y both be the set $\{a, b\}$. Let the topology for X be discrete and the topology for Y be indiscrete, consisting only of $\varnothing$ and the whole space. Every function $f : X \to Y$ is continuous. Suppose that $f(a) = f(b) = a$. Take $x_0 = b$ and $x_\alpha = b$ for all α. Then $\{f(x_\alpha)\}$ converges to both a and b. Hence we cannot evaluate $f(x_0)$ as just any limit of $\{f(x_\alpha)\}$; we have to pick the right limit.

PROOF. Suppose that f is continuous at x_0 and that $\{x_\alpha\}$ is a convergent net in X with limit x_0. Let V be any open neighborhood of $f(x_0)$. By continuity, there exists an open neighborhood U of x_0 such that $f(U) \subseteq V$. Since $x_\alpha \to x_0$, the members x_α of the net are eventually in U. Then $f(x_\alpha)$ is in $f(U) \subseteq V$ for the same α's, hence eventually. Therefore $\{f(x_\alpha)\}$ converges to $f(x_0)$.

Conversely suppose that $x_\alpha \to x_0$ always implies $f(x_\alpha) \to f(x_0)$. We are to show that f is continuous. If V is an arbitrary open neighborhood of $f(x_0)$, we seek some open neighborhood of x_0 that maps into V under f. Assuming that there is no such neighborhood for some V, we can find, for each neighborhood U of x_0, some x_U in U such that $f(x_U)$ is not in V. Then $x_U \to x_0$, but $f(x_U)$ does not have limit $f(x_0)$ because $f(x_U)$ is never in V. This is a contradiction, and we conclude that some U maps into V under f; thus f is continuous.

Proposition 10.35. Let $X = \bigtimes_{s \in S} X_s$ be the product of topological spaces X_s, and let $p_s : X \to X_s$ be the s^{th} coordinate function. Then a net $\{x_\alpha\}$ in X converges to some x_0 in X if and only if the net $\{p_s(x_\alpha)\}$ in X_s converges to $p_s(x_0)$ for each s in S.

REMARK. This is the sense in which the product topology is the topology of pointwise convergence. In combination with Corollary 10.33, this proposition simplifies the problem of deciding when a subset of a product space is closed in the product topology.

PROOF. If $\{x_\alpha\}$ converges to x_0, then Proposition 10.34 and the continuity of p_s together imply that $\{p_s(x_\alpha)\}$ converges to $p_s(x_0)$.

Conversely suppose that $\{p_s(x_\alpha)\}$ converges to $p_s(x_0)$ for all s. Fix s. If U_s is an open neighborhood of $p_s(x_0)$ in X_s, then $\{p_s(x_\alpha)\}$ is eventually in U_s. Hence there is some α_0 such that $p_s(x_\alpha)$ is in U_s whenever $\alpha_0 \leq \alpha$. For the same values of α, $\{x_\alpha\}$ is in $p_s^{-1}(U_s)$. Thus $\{x_\alpha\}$ is eventually in $p_s^{-1}(U_s)$.

Any neighborhood N of x_0 in X contains some basic open neighborhood of the form $U = p_{s_1}^{-1}(U_{s_1}) \cap \cdots \cap p_{s_n}^{-1}(U_{s_n})$. It follows from the result of the previous paragraph that $\{x_\alpha\}$ is eventually in each $p_s^{-1}(U_s)$, hence is eventually in the intersection U, and hence is eventually in N. Therefore $\{x_\alpha\}$ converges to x_0.

One can express also the notion of compactness in terms of nets, the idea being that compactness of X is equivalent to the fact that every net in X has a convergent subnet, for an appropriate definition of "subnet." The remainder of this section will deal with this question. Carrying out the details of this equivalence is harder than what we have done so far with nets. Actually, the main benefit of the equivalence is the resulting simplification to proofs of compactness, especially to the proof of the Tychonoff Product Theorem. Since we have already proved the Tychonoff Product Theorem without nets, the material in the remainder of this section will be used only in minor ways in the rest of the book.[6]

Let D and E be directed sets. A function from E to D, written $\mu \mapsto \alpha_\mu$, is **cofinal**[7] if for any β in D, there is a ν in E such that $\beta \leq \alpha_\mu$ whenever $\nu \leq \mu$. If $\mu \mapsto \alpha_\mu$ is cofinal and if $\alpha \mapsto x_\alpha$ is a net from D to X, then the composition $\mu \mapsto x_{\alpha_\mu}$ is a net from E to X and is called a **subnet** of the net $\alpha \mapsto x_\alpha$.

The prototype of a subnet is a subsequence. In this case, D and E are both the set of positive integers, and the function from E to D is $k \mapsto n_k$. If the sequence is $\{a_n\}$, then the subnet/subsequence is $\{a_{n_k}\}$. For a general subnet one might expect that it would suffice always to take E to be a subset of D and to let the function from E to D be inclusion. However, this definition of subnet is insufficient to prove the desired characterization of compactness in terms of nets and subnets.

A net from a directed set D to a nonempty set X is called **universal** if for any subset A of X, the net is eventually in A or eventually in A^c. It of course cannot be eventually in both, since otherwise it would eventually be in the intersection, namely the empty set.

Proposition 10.36. Each net in a nonempty set X has a universal subnet.

REMARK. The proof will use Zorn's Lemma. Apart from this one use, the only other uses of the Axiom of Choice in the remainder of this section are transparent ones.

[6]Nets play a more significant role in the companion volume *Advanced Real Analysis*.

[7]This definition is not the standard one given in Kelley's *General Topology*, but it leads to the standard definition of "subnet."

PROOF. Let D be a directed set, and let $\alpha \mapsto x_\alpha$ be a net from D to X. Consider all families $\mathcal{C}_\beta$ of subsets of X that are closed under finite intersections and have the property, for each A in $\mathcal{C}_\beta$, that the net is frequently in A. There exists such a family, the singleton family $\{X\}$ being one. Partially order the set of such families by inclusion upward, saying that $\mathcal{C}_\beta \leq \mathcal{C}_{\beta'}$ when $\mathcal{C}_\beta \subseteq \mathcal{C}_{\beta'}$. In any chain of $\mathcal{C}_\beta$'s, let $\mathcal{C}_\gamma$ be the union of the sets in the various members of the chain. Since closure under intersection depends only on two sets at a time and since the other property of a $\mathcal{C}_\beta$ depends only on one set at a time, $\mathcal{C}_\gamma$ is again a family of this kind. By Zorn's Lemma let $\mathcal{C}$ be a maximal such family.

Let us prove for each subset A of X that either A or A^c is in $\mathcal{C}$. In fact, if for every B in $\mathcal{C}$, the net is frequently in $A \cap B$, then $\mathcal{C} \cup \{A\}$ is a family containing $\mathcal{C}$ and satisfying the two defining properties of one of our families. By maximality, $\mathcal{C} \cup \{A\} = \mathcal{C}$. Hence A is in $\mathcal{C}$. Assuming that A is not in $\mathcal{C}$, we obtain a set B in $\mathcal{C}$ such that the net fails to be frequently in $A \cap B$. Then B is a member of $\mathcal{C}$ such that the net is eventually in $(A \cap B)^c$.

Similarly if we assume that A^c is not in $\mathcal{C}$, we obtain a set B' in $\mathcal{C}$ such that the net is eventually in $(A^c \cap B')^c$. If neither A nor A^c is in $\mathcal{C}$, then the net is eventually in

$$\begin{aligned}(A \cap B)^c \cap (A^c \cap B')^c &= (A^c \cup B^c) \cap (A \cup B'^c)\\ &= (A^c \cap (A \cup B'^c)) \cup (B^c \cap (A \cup B'^c))\\ &= (A^c \cap B'^c) \cup (B^c \cap (A \cup B'^c))\\ &\subseteq B'^c \cup B^c = (B \cap B')^c,\end{aligned}$$

and it cannot be frequently in $B \cap B'$. This contradicts the fact that $B \cap B'$ is in $\mathcal{C}$ because $\mathcal{C}$ is closed under finite intersections. This completes the proof that either A or A^c has to be in $\mathcal{C}$.

The members of $\mathcal{C}$ form a directed set under inclusion downward, i.e., with partial ordering $A \leq B$ if $A \supseteq B$. Form $\mathcal{C} \times D$ as a directed set under the definition in Example 4 at the beginning of this section. We construct a subnet as follows. For each ordered pair (A, β) in $\mathcal{C} \times D$, let $\alpha_{(A,\beta)}$ be an element of D with $\beta \leq \alpha_{(A,\beta)}$ and with $x_{\alpha_{(A,\beta)}}$ in A; this choice is possible since D is directed and the given net is frequently in A. The function $(A, \beta) \mapsto \alpha_{(A,\beta)}$ is cofinal because for any $\beta \in D$, the domain value (A, β) has $\beta \leq \alpha_{(B,\gamma)}$ whenever $(A, \beta) \leq (B, \gamma)$. Thus $(A, \beta) \mapsto x_{\alpha(A,\beta)}$ is a subnet.

To complete the proof, we show that this subnet is universal. For any subset A of X, we have seen that either A or A^c has to be in $\mathcal{C}$. Without loss of generality, assume that A is in $\mathcal{C}$. For any fixed β, the inequality $(A, \beta) \leq (B, \gamma)$ implies that $x_{\alpha_{(B,\gamma)}}$ is in the subset B of A, and hence the subnet is eventually in A.

Proposition 10.37. The following three statements about a topological space X are equivalent:

(a) X is compact,
(b) every universal net in X is convergent,
(c) every net in X has a convergent subnet.

PROOF. To prove that (a) implies (b), let $\{x_\alpha\}$ be a universal net in X, and suppose that $\{x_\alpha\}$ is not convergent. For each x in X, there is then an open neighborhood U_x of x such that $\{x_\alpha\}$ is not eventually in U_x. Since the net is universal, it is eventually in $(U_x)^c$ for each x. The open sets U_x cover X. By compactness, let $\{U_{x_1}, \dots, U_{x_n}\}$ be a finite subcover. The net is eventually in each $(U_{x_j})^c$ and hence is eventually in their intersection. But their intersection is empty since $X = \bigcup_{j=1}^n U_{x_j}$. We have arrived at a contradiction, and thus $\{x_\alpha\}$ must be convergent.

Statement (b) implies statement (c) since every net has a universal subnet, by Proposition 10.36.

To prove that (c) implies (a), suppose that X is noncompact. We shall produce a net with no convergent subnet. If $\mathcal{U}$ is an open cover of X with no finite subcover, we shall use $\mathcal{U}$ to define a directed set. Let $\mathcal{F}$ be the set of all finite subcollections of members of $\mathcal{U}$. This is directed under inclusion upward: $\alpha \leq \beta$ if $\alpha \subseteq \beta$. For each α in $\mathcal{F}$, the set $X - \bigcup_{U\in\alpha} U$ is not empty since $\mathcal{U}$ has no finite subcover, and we let x_α be an element of $X - \bigcup_{U\in\alpha} U$. Then $\alpha \mapsto x_\alpha$ is a net. Suppose that $\{x_\alpha\}$ has a convergent subnet, with some x_0 as limit. For any neighborhood N of x_0, $\{x_\alpha\}$ is frequently in N. Since $\mathcal{U}$ is a covering, there is some U in $\mathcal{U}$ with x_0 in U. By construction, $\{x_\alpha\}$ is not in U as soon as α has $\{U\} \leq \alpha$. We conclude that no subnet of $\{x_\alpha\}$ converges.

Proposition 10.37 gives the statement about general topological spaces that extends the equivalence of the Bolzano–Weierstrass property and the Heine–Borel property of closed bounded subsets of Euclidean space. To illustrate the power of nets, we can now use them to give a second proof of the Tychonoff Product Theorem (Theorem 10.27).

SECOND PROOF OF TYCHONOFF PRODUCT THEOREM. Let $X = \bigtimes_{s\in S} X_s$ be given with each X_s compact, let $p_s : X \to X_s$ be the s^{th} coordinate function, and let $\{x_\alpha\}$ be a universal net in X. Fix s, and let A_s be any subset of X_s. Since the net is universal, it is eventually in $p_s^{-1}(A_s)$ or in $(p_s^{-1}(A_s))^c$. Since $(p_s^{-1}(A_s))^c = p_s^{-1}((A_s)^c)$, the net $\{p_s(x_\alpha)\}$ is eventually in A_s or in $(A_s)^c$. Thus $\{p_s(x_\alpha)\}$ is a universal net in X_s. By Proposition 10.37 and the compactness of X_s, $\{p_s(x_\alpha)\}$ converges to some member x_s of X_s. Now let s vary. Forming the member x of X with $p_s(x) = x_s$ for all s and applying Proposition 10.35, we see that $x_\alpha \to x$. By Proposition 10.37, X is compact.

6. Quotient Spaces

If X is a topological space and $\sim$ is an equivalence relation on X, then we saw in Section 1 that the set $X/\sim$ of equivalence classes inherits a natural topology known as the "quotient topology." If $q : X \to X/\sim$ is the **quotient map**, then a subset U of $X/\sim$ is defined to be open in the **quotient topology** if $q^{-1}(U)$ is open in X. The quotient topology is then the finest topology on $X/\sim$ that makes the quotient map continuous.

Without some assumption that relates the equivalence relation to the topology of X, we cannot expect much from general quotient spaces. In this section we shall investigate situations in which the quotient space does have reasonable properties. Ultimately our interest will be in four situations, some of which are hinted at in Section 1:

(i) the passage from a regular topological space to the quotient when the equivalence relation is that $x \sim y$ if x is in $\{y\}^{\mathrm{cl}}$ (Proposition 10.7),

(ii) the passage from a compact Hausdorff space X to the quotient when the equivalence relation is closed as a subset of $X \times X$ (to be discussed in Problem 11 at the end of the chapter),

(iii) the passage from a "topological vector space" or "topological group" to a coset space (to be discussed in the companion volume *Advanced Real Analysis*),

(iv) the piecing together of a "manifold," or a "vector bundle," or a "covering space" from local data (to be discussed in the companion volume *Advanced Real Analysis*).

We begin with some general facts. The first is a kind of "universal mapping property" for all quotient spaces. Its corollary describes a situation in which we can recognize a given space as a quotient even if it was not constructed that way: we say that a function $F : X \to Y$ is **open** if F carries open sets to open sets.

Proposition 10.38.

(a) Let $F : X \to Y$ be a continuous function between topological spaces, let $\sim$ be an equivalence relation on X, and let $q : X \to X/\sim$ be the quotient map. Suppose that F has the property that $F(x_1) = F(x_2)$ whenever $x_1 \sim x_2$, so that there exists a well-defined function $f : X/\sim \to Y$ such that $F = f \circ q$. Then f is continuous.

(b) The quotient $X/\sim$ is characterized by the property in (a) in the following sense. Suppose that $q' : X \to Z$ is any continuous function of X onto a topological space Z such that

(i) $x_1 \sim x_2$ implies $q'(x_1) = q'(x_2)$,

(ii) whenever $F : X \to Y$ is a continuous function such that $x_1 \sim x_2$

implies $F(x_1) = F(x_2)$, there exists a continuous function $f' : Z \to Y$ with $F = f' \circ q'$.

Then Z is canonically homeomorphic to $X/\sim$.

PROOF. In (a), we want to know that $f^{-1}(U)$ is open in $X/\sim$ whenever U is open in Y. By definition of the quotient topology, $f^{-1}(U)$ is open in $X/\sim$ if and only if $q^{-1}(f^{-1}(U))$ is open in X. This set is $F^{-1}(U)$, which is open since F is assumed continuous.

In (b), suppose Z and q' are such that $q' : X \to Z$ has the stated properties. We apply the result of (a) with F taken to be $q : X \to X/\sim$, and the property of Z gives us a continuous function $f' : Z \to X/\sim$ such that $q = f' \circ q'$. Then we apply the result of (a) with F taken to be $q' : X \to Z$, and (a) shows that the function $f : X/\sim \to Z$ with $q' = f \circ q$ is continuous. Combining these two equations gives us $q = f' \circ f \circ q$ and $q' = f \circ f' \circ q'$. Thus $f' \circ f$ is the identity on the image of q, and $f \circ f'$ is the identity on the image of q'. Since q is onto $X/\sim$ and q' is onto Z, $f : X/\sim \to Z$ is a homeomorphism.

Corollary 10.39. Let $F : X \to Y$ be a continuous function from one topological space onto another, and define $x_1 \sim x_2$ if $F(x_1) = F(x_2)$. Let $q : X \to X/\sim$ be the quotient map, and let $f : X/\sim \to Y$ be the continuous map such that $F = f \circ q$. If F is open, then f is a homeomorphism and hence Y can be regarded as a quotient of X.

REMARK. The continuity of f is the conclusion of Proposition 10.38a.

PROOF. The function $f : X/\sim \to Y$ is continuous, one-one, and onto. To see that f is open and hence is a homeomorphism, let an open set U in $X/\sim$ be given. Then $F(q^{-1}(U))$ is open because q is continuous and F is open. Since $F(q^{-1}(U)) = f(q(q^{-1}(U))) = f(U)$, we see that $f(U)$ is open. Hence f is open.

EXAMPLE. Let $X = \bigtimes_{s \in S} X_s$ be a product of topological spaces, fix s in S, and let $p_s : X \to X_s$ be the s^{th} coordinate function. We shall show that p_s is open, so that X_s can be regarded as the quotient of X by the relation that $x_1 \sim x_2$ if $p_{s'}(x_1) = p_{s'}(x_2)$ for all $s' \neq s$. If U is an open set in X and x is in U, then we can find a basic open set $V_x = p_{s_1}^{-1}(U_1) \cap \cdots \cap p_{s_n}^{-1}(U_n)$ about x that is contained in U. Then $p_s(V_x)$ equals U_j if $s = s_j$, and it equals X_s if s is not equal to any s_j. In either case, $p_s(V_x)$ is open. Thus $p_s(U)$ contains a neighborhood of each of its points and must be an open set. So p_s is open.

A key desirable property of a quotient space is that it is Hausdorff. The Hausdorff property is what makes limits unique, after all, and it therefore paves the way to doing some analysis with the space. The next proposition gives a useful necessary condition and a useful sufficient condition.

Proposition 10.40. Let X be a topological space, let $\sim$ be an equivalence relation on X, and let R be the subset $\{(x_1, x_2) \mid x_1 \sim x_2\}$ of $X \times X$. If the quotient topology on $X/\sim$ is Hausdorff, then R is a closed subset of $X \times X$. Conversely if R is a closed subset of $X \times X$ and if the quotient map $q : X \to X/\sim$ is open, then $X/\sim$ is Hausdorff.

PROOF. Suppose that $X/\sim$ is Hausdorff. If (x, y) is not in R, then $q(x)$ and $q(y)$ are distinct points in $X/\sim$. Find disjoint open sets U and V in $X/\sim$ such that $q(x)$ is in U and $q(y)$ is in V. Then $q^{-1}(U)$ and $q^{-1}(V)$ are open sets in X with the property that no member of $q^{-1}(U)$ is equivalent to any member of $q^{-1}(V)$. Thus $q^{-1}(U) \times q^{-1}(V)$ is an open neighborhood of (x, y) that does not meet R. Hence R is closed.

Conversely if R is closed and (x, y) is not in R, then there exists a basic open set $U \times V$ of $X \times X$ containing (x, y) that does not meet R. The sets $q(U)$ and $q(V)$ are open in $X/\sim$ since q is open, they are disjoint since no member of U is equivalent to a member of V, and they are neighborhoods of $q(x)$ and $q(y)$, respectively. Thus $X/\sim$ is Hausdorff.

A special case is the situation with a pseudometric space in which the equivalence relation is that $x \sim y$ if x and y are at distance 0 from one another. A generalization of this relation was given in Proposition 10.7, which said that in a regular topological space the relation $x \sim y$ if x is in $\{y\}^{\mathrm{cl}}$ is an equivalence relation. The corollary to follow gives properties of the quotient space when this equivalence relation is used.

Corollary 10.41. Let X be a regular topological space, let $\sim$ be the equivalence relation defined by saying that $x \sim y$ if x is in $\{y\}^{\mathrm{cl}}$, and let $q : X \to X/\sim$ be the quotient map. Then

(a) q is open, and every open set in X is the union of equivalence classes,
(b) $X/\sim$ is regular and Hausdorff,
(c) X normal implies $X/\sim$ normal,
(d) X separable implies $X/\sim$ separable.

PROOF. First we show that every open set is a union of equivalence classes. Suppose that x is in an open set U in X. Let $x \sim y$. If y were not in U, then y would be in the closed set U^c and hence $\{y\}^{\mathrm{cl}}$ would be contained in U^c. Since $x \sim y$, x is in $\{y\}^{\mathrm{cl}}$, and we are led to the contradiction that x would be in U^c, hence in $U \cap U^c = \varnothing$. So U is a union of equivalence classes. Then it follows that $q^{-1}(q(U)) = U$, and the set $q(U)$ has the property that its inverse image is open in X. By definition of the quotient topology, $q(U)$ is open. Therefore q is an open map. This proves (a).

To prove the Hausdorff property in (b), we shall apply Proposition 10.40. Since (a) shows that q is open, it is enough to show that the subset $R = \{(x, y) \mid x \sim y\}$ of $X \times X$ is closed. If (x, y) is not in R, then x is not in $\{y\}^{\mathrm{cl}}$. By regularity of X, choose disjoint open sets U and V in X such that x is in U and $\{y\}^{\mathrm{cl}} \subseteq V$. Since U and V are unions of equivalence classes and are disjoint, no member of U is equivalent to any member of V. Therefore $(U \times V) \cap R = \varnothing$, and every point of R^c has an open neighborhood lying in R^c. Hence R is closed.

As a result of (a), the open sets in X are in one-one correspondence via q with the open sets in $X/\sim$, and the same thing is true for the closed sets. Under this correspondence disjoint sets correspond to disjoint sets. Then regularity in (b), as well as conclusions (c) and (d), follow immediately.

7. Urysohn's Lemma

According to Proposition 10.31, a Hausdorff topological space has unique limits for convergent sequences and nets. Corollary 10.41 shows that regularity of a space makes it possible to pass to a natural quotient space that is regular and Hausdorff. The following theorem exhibits a special role for the condition that a space be normal.

Theorem 10.42. (Urysohn's Lemma). If E and F are disjoint closed sets in a normal topological space X, then there exists a continuous function f from X into $[0, 1]$ that is 0 on E and is 1 on F.

PROOF. Proposition 10.5c shows in a normal space that between a closed set and a larger open set we can always interpolate an open set and its closure. Starting from $E \subseteq F^c$, we find an open set $U_{1/2}$ with

$$E \subseteq U_{1/2} \subseteq (U_{1/2})^{\mathrm{cl}} \subseteq F^c.$$

Then we can find open sets $U_{1/4}$ and $U_{3/4}$ with

$$E \subseteq U_{1/4} \subseteq (U_{1/4})^{\mathrm{cl}} \subseteq U_{1/2} \subseteq (U_{1/2})^{\mathrm{cl}} \subseteq U_{3/4} \subseteq (U_{3/4})^{\mathrm{cl}} \subseteq F^c.$$

Proceeding inductively on n, we obtain, for each diadic rational number $r = m/2^n$ with $0 < r < 1$, an open set U_r between E and F^c such that $r < s$ implies $(U_r)^{\mathrm{cl}} \subseteq U_s$. Put $U_1 = X$. For each x in X, define $f(x)$ to be the greatest lower bound of all r such that x is in U_r. Then f is 0 on E, is 1 on F, and has values in $[0, 1]$. To see that f is continuous, let x be given, let r and s be diadic rationals in $(0, 1)$ with $r < f(x) < s$, and choose diadic rationals r' and s' with $r < r' < f(x) < s' < s$. (If $f(x) = 0$, we omit r and r'; if $f(x) = 1$, we omit s and s'.) We are to produce an open neighborhood U of x with $f(U) \subseteq (r, s)$. If $U = U_{s'} - (U_{r'})^{\mathrm{cl}}$, then U is open with $r' \leq f(U) \leq s'$. Thus $r < f(U) < s$ as required. We conclude that f is continuous.

EXAMPLE. In Example 4 of Section 2, we produced a certain Hausdorff regular space X that is not normal, but we deferred the proof that X is not normal until we had Urysohn's Lemma in hand. We can now give that missing proof. As a set, X is the closed upper half plane $\{\operatorname{Im} z \geq 0\}$ in $\mathbb{C}$. A base for the topology in question consists of all open disks in X that do not meet the x axis, together with all open disks in X that are tangent to the x axis; the latter sets are to include the point of tangency. For a point p on the x axis, the open disks of rational radii with point of tangency p form a countable local base. Arguing by contradiction, suppose that X is normal. Any subset of the x axis in X is closed in X, and we take E to be the set of rationals on the axis and F to be the set of irrationals on the axis. Urysohn's Lemma (Theorem 10.42) supplies a continuous function $f : X \to [0, 1]$ such that $f(E) = 0$ and $f(F) = 1$. Define a sequence of functions $f_n : \mathbb{R} \to [0, 1]$ by $f_n(x) = f\big(x, \frac{1}{n}\big)$, the notation (x, y) indicating a point in the (x, y) plane. The functions f_n are continuous in the ordinary topology on $\mathbb{R}$ since the topology on X is the ordinary topology of the half plane as long as we stay away from the x axis. At any point $(x, 0)$ of the x axis, the sets

$$U_m = \{x, 0\} \cup B\big(\tfrac{1}{m}; \big(x, \tfrac{1}{m}\big)\big)$$

form a local base at $(x, 0)$, and $\big(x, \frac{1}{n}\big)$ is in U_m for $n \geq m$. The continuity of f therefore yields $\lim_n f\big(x, \frac{1}{n}\big) = f(x, 0)$. In other words, $\lim_n f_n$ exists pointwise on $\mathbb{R}$ and equals the indicator function of the set of irrationals. The sequence $\{f_n\}$ is therefore a sequence of continuous real-valued functions on $\mathbb{R}$ whose pointwise limit is everywhere discontinuous. However, Theorem 2.54 implies that the set of discontinuities of the limit function is of first category in $\mathbb{R}$, and the Baire Category Theorem (Theorem 2.53) implies that $\mathbb{R}$ is not of first category in itself. Thus we have a contradiction, and we conclude that X cannot be normal.

Corollary 10.43. If E and F are disjoint closed sets in a compact Hausdorff space X, then there exists a continuous function $f : X \to [0, 1]$ that is 0 on E and is 1 on F.

PROOF. This follows by combining Proposition 10.16 and Theorem 10.42.

Corollary 10.44. If K and F are disjoint closed sets in a locally compact Hausdorff space X and if K is compact, then there exists a continuous function $f : X \to [0, 1]$ that is 1 on K, is 0 on F, and has compact support.

PROOF. Using Proposition 10.19, regard X as an open subset of the one-point compactification X^*. Proposition 10.20 shows that the compact space X^* is Hausdorff. Choose disjoint open sets U and V in X by Corollary 10.22 such that $K \subseteq U$ and $F \subseteq V$. Choose L compact in X by Corollary 10.23 such that $K \subseteq L^o$. Then $M = L \cap (X - V)$ is compact in X by Proposition 10.14, and $K \subseteq L^o \cap U \subseteq L^o \cap (X - V)^o \subseteq (L \cap (X - V))^o = M^o$. Hence K and $X^* - M^o$ are disjoint compact sets in X^*. Corollary 10.43 produces a continuous $g : X^* \to [0, 1]$ such that g is 1 on K and is 0 on $X^* - M^o$. Since $F \subseteq V \subseteq (X - L) \cup V = X - (L \cap (X - V)) = X - M \subseteq X - M^o \subseteq X^* - M^o$, the function $f = g\big|_X$ has the required properties.

8. Metrization in the Separable Case

A problem about topological spaces, now completely solved, is to characterize those topologies that arise from metric spaces. Such a space is said to be **metrizable**. We consider only the separable case and prove the following theorem.

Theorem 10.45 (Urysohn Metrization Theorem). Any separable regular Hausdorff space X is homeomorphic to a subspace of the Hilbert cube $C = \bigtimes_{n=1}^{\infty} [0, 1]$ and is therefore metrizable.

PROOF. The Hilbert cube C is seen as a metric space in Example 11 in Section II.1, Corollary 10.29 identifies it as a product space, and the Tychonoff Product Theorem (Theorem 10.27) shows that it is compact. Let $p_n : X \to [0, 1]$ be the n^{th} coordinate function.

By Corollary 10.10, X is normal. Fix a countable base $\mathcal{B}$ for the open sets. Enumerate the countable set of pairs (U, V) of members of $\mathcal{B}$ such that $U^{\text{cl}} \subseteq V$. To the n^{th} pair, associate by Urysohn's Lemma (Theorem 10.42) a continuous function $f_n : X \to [0, 1]$ such that f_n is 1 on U^{cl} and is 0 on V^c. Let $F : X \to C$ be defined by "$F(x)$ is the sequence whose n^{th} term is $f_n(x)$." We are to show that F is continuous, is one-one, and is open as a function onto $F(X)$.

The continuity of $p_n \circ F = f_n$ for each n means that $F^{-1} p_n^{-1}$ of any open set in $[0, 1]$ is open in C. Since F^{-1} of a basic open set in C is the finite intersection of the various $F^{-1} p_n^{-1}$'s of open sets, F is continuous.

To see that F is one-one, let x and y be distinct points of x. By Proposition 10.6c, X Hausdorff implies that $\{y\}$ is closed and hence that $\{y\}^c$ is an open neighborhood of x. Choose a basic open set V containing x and contained in $\{y\}^c$. By Proposition 10.5b and the regularity of X, choose a basic open set U containing x such that $U^{\text{cl}} \subseteq V$. Then (U, V) is one of our pairs, and the corresponding function f_n has $f_n(x) = 1$ and $f_n(y) = 0$. Hence $F(x) \neq F(y)$, and F is one-one.

To see that F carries open sets of X to open sets in $F(X)$, let W be open in X, and fix x in W. Arguing as in the previous paragraph, we can find basic open sets U and V such that x is in U and $U^{\text{cl}} \subseteq V \subseteq W$. The corresponding f_n then has $f_n(x) = 1$ and $f_n(V^c) = 0$. Hence $f_n(W^c) = 0$. The set N_x of y's such that $f_n(y) > 0$ is open in X and contains x. The product $(0, 1]_n \times \big(\bigtimes_{k \neq n} [0, 1]_k\big)$ is open in C, and its intersection with $F(X)$ is the same as $F(N_x) \cap F(X)$. Thus $F(N_x) \cap F(X)$ is relatively open in $F(X)$. Then $F(x)$ lies in this relatively open set, which in turn lies in $F(W)$, and it follows that $F(W)$ is a relatively open neighborhood of each of its members.

Corollary 10.46. Every separable compact Hausdorff space is metrizable.

PROOF. This is immediate from Proposition 10.16 and Theorem 10.45.

9. Ascoli–Arzelà and Stone–Weierstrass Theorems

In Section II.10 we studied Ascoli's Theorem (Theorem 2.56) and the Stone–Weierstrass Theorem (Theorem 2.58) as tools for working with continuous functions on compact metric spaces. In turn, these theorems were illuminating generalizations of results about continuous functions on closed bounded intervals of the line, particularly the classical version of Ascoli's Theorem (Theorem 1.22) and the Weierstrass Approximation Theorem (Theorem 1.52). In this section we shall extend these results to the setting of continuous functions on compact Hausdorff spaces. The proof of the extended Ascoli theorem will be our first example of how the Cantor diagonal process gets replaced by an application of the Tychonoff Product Theorem (Theorem 10.27) when one is dealing with an uncountable number of limiting situations at once. The Stone–Weierstrass Theorem in the more general setting becomes in part a tool for dealing with large abstract compact Hausdorff spaces that arise in functional analysis. The starting point for this investigation is the general form of Alaoglu's Theorem,[8] which says that the closed unit ball in the dual X^* of a normed linear space X is compact in the weak-star topology; closed subsets of this space play a foundational role in the theory of Banach algebras.

We work in this section with a compact Hausdorff space X and with the algebra $C(X)$ of bounded continuous scalar-valued functions on X. The scalars may be real or complex. Corollary 10.13 shows that if f is a continuous scalar-valued function on X, then $|f|$ attains its maximum value on X. The set $C(X)$ is a subspace of the normed linear space $B(X)$ of bounded scalar-valued functions on X, the norm being $\|f\|_{\sup} = \sup_{x\in X} |f(x)|$. Convergence in $B(X)$ is uniform convergence. Proposition 10.30 shows that $C(X)$ is a closed subspace of $B(X)$ and is complete as a metric space.

We begin with the extended Ascoli theorem. Let $\mathcal{F} = \{f_\alpha\}$ be a set of scalar-valued functions on the compact Hausdorff space X. We say that $\mathcal{F}$ is **equicontinuous** at x in X if for each $\epsilon > 0$, there is an open neighborhood $U_{x,\epsilon}$

[8] A preliminary form of this theorem was given as Theorem 5.58. The general form appears in the companion volume *Advanced Real Analysis*.

of x such that $|f_\alpha(y) - f_\alpha(x)| < \epsilon$ for all y in $U_{x,\epsilon}$ and all f_α in $\mathcal{F}$. We say that $\mathcal{F}$ is **equicontinuous** if it is equicontinuous at each point. Not having a metric to compare different points of X, we no longer define a notion of "uniform equicontinuity."

It is immediate from the definition that any subset of an equicontinuous family is equicontinuous. The definition of equicontinuity at x reduces to the definition of continuity if $\mathcal{F}$ has just one member, and therefore every member of an equicontinuous family is continuous.

As in Section II.10 the set $\mathcal{F}$ is **uniformly bounded** on X if it is pointwise bounded at each $x \in X$ and if the bound for the values $|f(x)|$ with $f \in \mathcal{F}$ can be taken independent of x.

Lemma 10.47. If $\mathcal{F} = \{f_\alpha\}$ is equicontinuous at x in X, then the closure $\mathcal{F}^{\text{cl}}$ of $\mathcal{F}$ in the product topology on $\mathbb{C}^X$ is equicontinuous at x.

REMARK. Consequently every member of $\mathcal{F}^{\text{cl}}$ is continuous at x.

PROOF. Let $U_{x,\epsilon}$ be as in the definition of equicontinuity of $\mathcal{F}$ at x. For each $\epsilon > 0$, the set of functions $f \in \mathbb{C}^X$ such that

$$|f(y) - f(x)| \le \epsilon$$

for a particular y in X is a closed subset of $\mathbb{C}^X$. Thus the set of functions $f \in \mathbb{C}^X$ such that this inequality holds for all y in $U_{x,\epsilon}$, being an intersection of closed sets, is closed, and it contains $\mathcal{F}$. In turn, the intersection G of these sets taken over all $\epsilon > 0$ is closed in $\mathbb{C}^X$ and contains $\mathcal{F}$. For each $\epsilon > 0$, each g in this closure G satisfies the inequality $|g(y) - g(x)| < 2\epsilon$ whenever y is in $U_{x,\epsilon}$. Therefore G is equicontinuous at x, and so is its subset $\mathcal{F}^{\text{cl}}$.

Theorem 10.48 (Ascoli–Arzelà Theorem). If $\{f_n\}$ is an equicontinuous family of scalar-valued functions defined on a compact Hausdorff space X and if $\{f_n\}$ has the property that $\{f_n(x)\}$ is bounded for each x, then $\{f_n\}$ has a uniformly convergent subsequence.

PROOF. We may assume that there are infinitely many distinct functions f_n, since otherwise the assertion is trivial. Let $|f_n(x)| \le c_x$ for all n, and form the product space $C = \bigtimes_{x \in X} \{z \in \mathbb{C} \bigm| |z| \le c_x\}$. The space C is compact by the Tychonoff Product Theorem (Theorem 10.27), and we are now assuming that there are infinitely many members of the sequence $\{f_n\}$ in the space. Let S be the image of the sequence as a subset of C. If S were to contain all its limit points, then each f_n would have an open neighborhood in C disjoint from the rest of S; these open sets and S^c would form an open cover of C with no finite subcover, in contradiction to compactness of C. Thus S has a limit point f not in S. By Lemma 10.47 and the remarks before it, the family $S \cup \{f\}$ is equicontinuous.

Let $\epsilon > 0$. We shall complete the proof by producing an f_N in S such that $|f_N(x) - f(x)| < \epsilon$ for all x. By equicontinuity find an open neighborhood U_x for each x such that $y \in U_x$ implies

$$|f_n(y) - f_n(x)| < \epsilon/3 \qquad \text{for all } n$$

and

$$|f(y) - f(x)| < \epsilon/3.$$

The open sets U_x cover X, and finitely many of them suffice to cover, by the compactness of X. Thus there are finitely many points $x_1, \ldots, x_k$ in X with the property that for each y in X, there is some x_j with $1 \le j \le k$ such that

$$|f_n(y) - f_n(x_j)| < \epsilon/3 \qquad \text{and} \qquad |f(y) - f(x_j)| < \epsilon/3$$

for all n. Since f is a limit point of S, choose N such that

$$|f_N(x_j) - f(x_j)| < \epsilon/3 \qquad \text{for } 1 \le j \le k.$$

Then for every y in X, there is an x_j such that

$$|f_N(y) - f(y)| \le |f_N(y) - f_N(x_j)| + |f_N(x_j) - f(x_j)| + |f(x_j) - f(y)| < \epsilon.$$

Thus f_N is within distance ϵ of f, as asserted.

Corollary 10.49. If X is a compact Hausdorff space, then a subset $\mathcal{F} = \{f_\alpha\}$ of $C(X)$ is compact if and only if

(a) $\mathcal{F}$ is closed in $C(X)$,
(b) the set $\{f_\alpha\}$ is pointwise bounded at each point in X, and
(c) $\mathcal{F}$ is equicontinuous.

In this case, $\mathcal{F}$ is uniformly bounded.

PROOF. Suppose that the three conditions hold. Being a subset of $C(X)$, $\mathcal{F}$ is a metric space under the restriction of the metric. By Theorem 2.36, $\mathcal{F}$ will be compact if we prove that every sequence has a convergent subsequence. Because of (b) and (c), Theorem 10.48 shows that every sequence in $\mathcal{F}$ has a uniformly Cauchy subsequence. By (a) and the completeness of $C(X)$ given in Proposition 10.30, $\mathcal{F}$ is complete as a metric space. Hence the Cauchy subsequence converges to an element of $\mathcal{F}$.

Conversely suppose that $\mathcal{F}$ is compact. Property (a) follows since compact sets are closed in any metric space. For (b) and the stronger conclusion that $\mathcal{F}$ is uniformly bounded, the function $f \mapsto \|f\|_{\sup}$ is a continuous function on the compact set $\mathcal{F}$, and Corollary 10.13 shows that it is bounded. For the equicontinuity in (c), let $\epsilon > 0$ and x be given. Theorem 2.46 shows that $\mathcal{F}$ is totally bounded as a metric space. Hence we can find a finite set $f_1, \ldots, f_l$

in $\mathcal{F}$ such that each member f of $\mathcal{F}$ has $\sup_{y \in S} |f(y) - f_j(y)| < \epsilon$ for some j. By continuity of each f_i, choose an open neighborhood $U_{x,\epsilon}$ of x such that $|f_i(x) - f_i(y)| < \epsilon$ for $1 \leq i \leq l$ for all y in $U_{x,\epsilon}$. If f is some member of $\mathcal{F}$ and if f_j is the member of the finite set associated with f, then $y \in U_{x,\epsilon}$ implies

$$|f(y) - f(x)| \leq |f(y) - f_j(y)| + |f_j(y) - f_j(x)| + |f_j(x) - f(x)| < 3\epsilon.$$

Hence $\mathcal{F}$ is equicontinuous at each x in X.

Now we come to the extended Stone–Weierstrass Theorem. We are interested in showing that certain subalgebras of the algebra $C(X)$ of continuous scalar-valued functions on a compact Hausdorff space X are dense in $C(X)$. Except for the dropping of the assumption that X is metric, the assumptions and notation are the same as in Section II.10. In particular the scalars for the subalgebra and for $C(X)$ may be real or complex, and the statement of the theorem is slightly different in the two cases.

Theorem 10.50 (Stone–Weierstrass Theorem). Let X be a compact Hausdorff space.

(a) If $\mathcal{A}$ is a real subalgebra of real-valued members of $C(X)$ that separates points and contains the constant functions, then $\mathcal{A}$ is dense in the algebra of real-valued members of $C(X)$ in the uniform metric.

(b) If $\mathcal{A}$ is a complex subalgebra of $C(X)$ that separates points, contains the constant functions, and is closed under complex conjugation, then $\mathcal{A}$ is dense in $C(X)$ in the uniform metric.

REMARKS. Curiously, Urysohn's Lemma (Corollary 10.43) does not play a role in the proof. Instead, the role of Urysohn's Lemma is to ensure that $C(X)$ is large in applications, and then the present theorem has serious content. The actual proof of Theorem 10.50 is word-for-word the same as for Theorem 2.58, and there is no need to repeat it.

10. Problems

1. Let f and g be continuous functions from a topological space into a Hausdorff space Y.
   (a) Prove that the set of all points x in X for which $f(x) = g(x)$ is closed.
   (b) Prove that if $f(x) = g(x)$ for all x in a dense subset of X, then $f = g$.
2. **(Dini's Theorem)** Let X be a compact Hausdorff space. Suppose that the function $f_n : X \to \mathbb{R}$ is continuous, that $f_1 \leq f_2 \leq f_3 \leq \cdots$, and that $f(x) = \lim f_n(x)$ is continuous and is nowhere $+\infty$. Use the defining property of compactness to prove that $\{f_n\}$ converges to f uniformly on X.

3. **(Baire Category Theorem)** Prove that a locally compact Hausdorff space cannot be the countable union of closed nowhere dense sets.
4. Prove that a locally compact dense subset of a Hausdorff space is open.
5. This problem produces a locally compact Hausdorff space that is not normal. Verify the details of the construction. Let X be a countably infinite discrete space, and let Y be an uncountable discrete space. Let X^* and Y^* be their one-point compactifications, with the added points denoted by x_∞ and y_∞. The locally compact Hausdorff space is $Z = X^* \times Y^* - \{(x_\infty, y_\infty)\}$ with the relative topology. Two closed subsets that cannot be separated by disjoint open sets are $A = (\{x_\infty\} \times Y^*) - \{(x_\infty, y_\infty)\}$ and $B = (X^* \times \{y_\infty\}) - \{(x_\infty, y_\infty)\}$.
6. If X is compact, prove that each infinite subset of X has a limit point.
7. Let $\mathcal{U}$ be the family of subsets of $\mathbb{R}$ consisting of all sets $\{x \in \mathbb{R} \mid x < a\}$, together with $\varnothing$ and $\mathbb{R}$.
   (a) Prove that $\mathcal{U}$ is a topology for $\mathbb{R}$ and that it is not Hausdorff. (It is called the **upper topology** of $\mathbb{R}$.)
   (b) If $\{t_n\}_{n \in D}$ is a net in $\mathbb{R}$, define $\limsup_n t_n$ to be the infimum over n of $\sup_{m \in D,\ m \geq n}$. Prove that a net $\{t_n\}_{n \in D}$ in $\mathbb{R}$ converges to t relative to $\mathcal{U}$ if and only if $\limsup_n t_n \geq t$.
8. Let $(X, \mathcal{T})$ be a topological space, and let $\mathcal{U}$ be the upper topology of $\mathbb{R}$ as in the previous problem. A function $f : X \to \mathbb{R}$ is said to be **upper semicontinuous** if it is continuous with respect to $\mathcal{T}$ and $\mathcal{U}$.
   (a) Prove that upper semicontinuity of $f : X \to \mathbb{R}$ is equivalent to the condition that $\limsup f(x_n) \leq f(x)$ whenever $x_n \to x$ in X.
   (b) Prove that the function $f : \mathbb{R} \to \mathbb{R}$ that is 1 at $x = 0$ and is 0 elsewhere is upper semicontinuous.
   (c) Prove that if f and g are upper semicontinuous functions on X and if c is nonnegative real, then $f + g$ and cf are upper semicontinuous.
   (d) Prove that if $\{f_s\}_{s \in S}$ is a nonempty set of upper semicontinuous functions on X such that $\inf_{s \in S} f(x) > -\infty$ for all $x \in X$, then $\inf_{s \in S} f_s$ is upper semicontinuous.
   (e) Prove that if f is a bounded real-valued function on X, then there exists a unique smallest upper semicontinuous function f^- with $f^-(x) \geq f(x)$ for all x.
9. Let $(X, \mathcal{T})$ be a topological space. A function $f : X \to \mathbb{R}$ is **lower semicontinuous** if $-f$ is upper semicontinuous. In this case if f is bounded, let $f_- = -(-f)^-$, with the right side defined as in the previous problem. Let the **oscillation** Q_f of f be defined by $Q_f(x) = f^-(x) - f_-(x)$ for x in X.
   (a) Why is Q_f upper semicontinuous?
   (b) Prove that this definition agrees with the one in Section II.9.
   (c) Prove that f is continuous if and only if Q_f is identically 0.

10. Let X be a Hausdorff topological space in which there are two disjoint nonempty closed sets A and B. Let $\sim$ be the equivalence relation that identifies all elements of A with each other, identifies all elements of B with each other, and otherwise identifies no distinct points of X.
    (a) Prove that the subset of pairs (x, y) in $X \times X$ with $x \sim y$ is closed.
    (b) Give an example of this kind in which $X/\sim$ is not Hausdorff.
11. Let X be a compact Hausdorff space, and let $\sim$ be an equivalence relation on X such that the subset $R \subseteq X \times X$ of pairs (x, y) with $x \sim y$ is closed. Let $q : X \to X/\sim$ be the quotient map.
    (a) Prove for each $x \in X$ that $q^{-1}q(x)$ is a closed subset of X.
    (b) If $U \subseteq X$ is open, prove that $V = \{x \in X \mid q^{-1}q(x) \subseteq U\}$ is open by first proving that $V^c = p_2((U^c \times X) \cap R)$, where $p_2 : X \times X \to X$ is the projection to the second coordinate.
    (c) Prove that the compact quotient $X/\sim$ is Hausdorff.
    (d) Prove that the quotient map is **closed**, i.e., that closed sets map to closed sets.
    (e) Is the quotient map necessarily open?
    (f) As in one of the examples in Section 1, let X be the interval $[-\pi, \pi]$, and let S^1 be the unit circle in $\mathbb{C}$. Let $\sim$ be the equivalence relation that lets $-\pi$ and π be the only nontrivial pair of elements of X that are equivalent, and form $X/\sim$. Prove that $X/\sim$ is homeomorphic to S^1 and that under this identification the quotient map may be taken to be the function $p : X \to S^1$ given by $p(x) = e^{ix}$.

Problems 12–15 concern connectedness and connected components. Most of the definitions and proofs in the first three are rather similar to those in Chapter II (§II.8 and Problems 11–13) for the special case of metric spaces. A topological space X is **connected** if X cannot be written as $X = U \cup V$ with U and V open, disjoint, and nonempty. A subset E of X is **connected** if E is connected as a subspace of X, i.e., if E cannot be written as a disjoint union $(E \cap U) \cup (E \cap V)$ with U and V open in X and with $E \cap U$ and $E \cap V$ both nonempty.

12. (a) Prove that a continuous function between topological spaces carries connected sets to connected sets.
    (b) A **path** in a topological space X is a continuous function from a closed bounded interval $[a, b]$ into X. Why is the image of a path necessarily connected?
13. (a) If X is a topological space and $\{E_\alpha\}$ is a system of connected subsets of X with a point x_0 in common, prove that $\bigcup_\alpha E_\alpha$ is connected.
    (b) If X is a topological space and E is a connected subset of X, prove that the closure E^{cl} is connected.
14. (a) A topological space X is **pathwise connected** if for any two points x_1 and x_2 in X, there is some continuous $p : [a, b] \to X$ with $p(a) = x_1$ and $p(b) = x_2$. Why is a pathwise-connected space X necessarily connected?

(b) A topological space X is called **locally pathwise connected** if each point has arbitrarily small open neighborhoods that are pathwise connected. Prove that if X is connected and locally pathwise connected, then it is pathwise connected.

15. In a topological space X, define two points to be equivalent if they lie in a connected subset of X.
    (a) Show that this notion of equivalence is indeed an equivalence relation. The equivalence classes are called the **connected components** of X.
    (b) Prove that the connected components of X are closed sets.
    (c) Prove that the connected components of X are open sets if X is **locally connected**, i.e., if each point has arbitrarily small connected neighborhoods.

Problems 16–17 concern partitions of unity, which were introduced in Section III.5. An open cover $\mathcal{U}$ of a topological space is said to be **locally finite** if each point of x has a neighborhood that lies in only finitely many members of $\mathcal{U}$.

16. Suppose that $\mathcal{U}$ is a locally finite open cover of a normal space X. By applying Zorn's Lemma to the class of all functions F defined on subfamilies of $\mathcal{U}$ such that $F(U)$, for each U in the domain of F, is an open set with $F(U)^{\text{cl}} \subseteq U$ and
$$\Big(\bigcup_{U\in\text{domain}(F)} F(U)\Big) \cup \Big(\bigcup_{\substack{V\in\mathcal{U},\\ V\notin\text{domain}(F)}} V\Big) = X,$$
prove that it is possible to select, for each U in $\mathcal{U}$, an open set V_U such that $V_U^{\text{cl}} \subseteq U$ and such that $\{V_U \mid U \in \mathcal{U}\}$ is an open cover of X.

17. Prove that if $\mathcal{U}$ is a locally finite open cover of a normal space X, then it is possible to select, for each U in $\mathcal{U}$, a continuous function $f_U : X \to [0, 1]$ such that f_U is 0 outside U and such that $\sum_{U\in\mathcal{U}} f_U(x) = 1$ for all $x \in X$.

Problems 18–20 establish the Tietze Extension Theorem. Let X be a normal topological space, and let C be a closed subset of X. Suppose that f is a bounded real-valued continuous function defined on C. The theorem is that there exists a continuous function $F : X \to \mathbb{R}$ such that $F\big|_C = f$ and $\sup_{x\in X} |F(x)| = \sup_{x\in C} |f(x)|$.

18. Let $g_0 = f$, $c_0 = \sup_{x\in C} |g_0(x)|$, $P_0 = \{x \in C \mid g_0(x) \geq c_0/3\}$, and $N_0 = \{x \in C \mid g_0(x) \leq -c_0/3\}$. Show that there is a continuous function F_0 from X into $[-c_0/3, c_0/3]$ that is $c_0/3$ on P_0 and $-c_0/3$ on N_0.

19. In the previous problem, put $g_1 = g_0 - F_0$ on C, and let $c_1 = \sup_{x\in C} |g_1(x)|$. Show that $c_1 \leq \frac{2}{3}c_0$. When the result of the previous problem is applied to g_1 in order to produce a function F_1, what properties does F_1 have?

20. Show that iteration of the above results produces a sequence of continuous functions $F_n : X \to \mathbb{R}$ such that the series $\sum_{n=0}^{\infty} F_n(x)$ is uniformly convergent on X and such that the sum $F(x) = \sum_{n=0}^{\infty} F_n(x)$ is continuous. Show also that F has $F\big|_C = f$ and satisfies $\sup_{x\in X} |F(x)| = \sup_{x\in C} |f(x)|$.

Problems 21–28 concern order topologies. Suppose that X is a set with at least two elements and having a **total ordering**, i.e., a partial ordering $\leq$ such that

(i) $x \leq y$ and $y \leq x$ together imply $x = y$,
(ii) any x and y in the set have either $x \leq y$ or $y \leq x$.

Define $x < y$ to mean that $x \leq y$ and $x \neq y$. The **order topology** on X is the topology for which a base consists of all sets $\{x \mid x < b\}$, $\{x \mid a < x\}$, and $\{x \mid a < x < b\}$. For a nonempty subset Y of X, the terms "lower bound," "upper bound," "greatest lower bound," and "least upper bound" are defined in the expected way. Examples are given by the real line $\mathbb{R}$ with its usual topology, the set Ω of countable ordinals (as defined in Problems 25–33 at the end of Chapter V) with its order topology, and other examples given below.

21. Prove that every open interval $\{x \mid a < x < b\}$ in X is open and every closed interval $\{x \mid a \leq x \leq b\}$ is closed.
22. Prove that X is Hausdorff and regular in its order topology.
23. Prove that every nonempty subset with an upper bound has a least upper bound if and only if every every nonempty subset with a lower bound has a greatest lower bound. In this case, X is said to be **order complete**.
24. Suppose that X is order complete.
    (a) Prove that a nonempty subset Y of X is compact if and only if Y is closed and has a lower bound and an upper bound.
    (b) Prove that X is locally compact.
25. (a) Prove that if there exist a and b in X with $a < b$ and with no c such that $a < c < b$, then X is not connected, in the sense of Problems 12–15. Let us say that X has a **gap** when such a and b exist.
    (b) Prove that if X is order complete and has no gaps, then X is connected.
26. The set $X = [0, 1) \cup [2, 3)$ is totally ordered. Prove that this X is connected in its order topology, and conclude that the order topology is different from the relative topology for X as a subspace of $\mathbb{R}$.
27. The set $X = [0, 1) \cup (1, 2]$ is totally ordered. Prove that this X is not connected in its order topology but has no gaps.
28. Let X and Y be two totally ordered sets with at least two elements apiece. Define the **lexicographic ordering** on $X \times Y$ to be the total ordering given by $(x_1, y_1) \leq (x_2, y_2)$ if $x_1 < x_2$ or else $x_1 = x_2$ and $y_1 \leq y_2$.
    (a) Prove that the lexicographic ordering on $[0, 1] \times [0, 1]$ makes the space compact connected but not separable.
    (b) The **long line** is defined to be the product $\Omega \times [0, 1)$ with the lexicographic ordering, where Ω is the set of countable ordinals as defined in Problems 25–33 at the end of Chapter V. Prove that the long line is locally compact and connected but not separable.

CHAPTER XI

Integration on Locally Compact Spaces

Abstract. This chapter deals with the special features of measure theory when the setting is a locally compact Hausdorff space and when the measurable sets are the Borel sets, those generated by the compact sets.

Sections 1–2 establish the basic theorem, the Riesz Representation Theorem, which says that any positive linear functional on the space $C_{\text{com}}(X)$ of continuous scalar-valued functions of compact support on the underlying space X is given by integration with respect to a unique Borel measure having a property called regularity. The steps in the construction of the measure run completely parallel to those for Lebesgue measure if one regards the geometric information about lengths of intervals as being encoded in the Riemann integral. The Extension Theorem of Chapter V is the main technical tool.

Section 3 studies more closely the nature of regularity of Borel measures. One direct generalization of a Euclidean theorem is that the space of continuous functions of compact support in an open set is dense in every L^p space on that open set for $1 \le p < \infty$. A new result is the Helly–Bray Theorem—that any sequence of Borel measures of bounded total measure in a locally compact separable metric space has a weak-star convergent subsequence whose limit is a Borel measure.

Section 4 regards $C_{\text{com}}(X)$ as a normed linear space under the supremum norm and identifies the space of continuous linear functionals, with its norm, as a space of signed or complex Borel measures with a regularity property, the norm being the total-variation norm for the signed or complex Borel measure.

1. Setting

This chapter brings together the measure theory of Chapters V–VI and the theory of topological spaces of Chapter X in a way that takes many of our earlier most interesting examples into account. Specifically we shall study the special features of measure theory when the underlying space is a locally compact Hausdorff space. Our primary example from earlier is that of Lebesgue measure, first on $\mathbb{R}^1$ and then in $\mathbb{R}^N$. In $\mathbb{R}^1$ we considered also the class of all Stieltjes measures and showed how they are classified by monotone functions satisfying certain properties. We introduced Borel measures in $\mathbb{R}^N$ but did not attempt to classify them.

Along the way we saw glimpses of some other examples: The unit circle of $\mathbb{C}$ can be regarded as $[-\pi, \pi]$ if we identify $-\pi$ and π, and we obtained Lebesgue measure on the circle. As we saw, any open set or any compact set in $\mathbb{R}^N$ has a theory of Borel measures associated with it. Most of our concrete examples of such measures when $N > 1$ came about as a consequence of the change-of-variables formula for multiple integrals. Of particular interest is what we anticipated in Section VI.5 would ultimately come to be regarded as a "rotation-invariant measure on the sphere," the sphere S^{N-1} being a compact metric space. This measure corresponds to the expression $d\omega$ when Lebesgue measure dx on $\mathbb{R}^N$ is written in spherical coordinates and the factor $r^{N-1}\,dr$ is dropped. In the concrete case of $\mathbb{R}^3$, in which r is the radius, θ_1 is the latitude from the north pole, and θ_2 is the longitude, Lebesgue measure is given by $dx = r^2 \sin\theta_1\, d\theta_2\, d\theta_1\, dr$ and we have $d\omega = \sin\theta_1\, d\theta_2\, d\theta_1$. The change-of-variables formula in the N-variable case then reads

$$\int_{\mathbb{R}^N} f(x)\,dx = \int_{r=0}^{\infty}\int_{\omega\in S^{N-1}} f(r\omega)\,r^{N-1}\,d\omega\,dr$$

for every Borel measurable function $f \geq 0$ on $\mathbb{R}^N$. We shall be making sense of $d\omega$ as a genuine measure on S^{N-1} in the course of the present chapter.

In the opposite direction it is important not to get the idea that all important measure-theoretic examples in mathematics arise from locally compact Hausdorff spaces. Examples that arise from probability theory need not fit this pattern. This fact becomes clearer after one encounters some specific measure spaces that arise in the theory.[1]

Let us turn to the setting of this chapter, a locally compact Hausdorff space X. In order that the measure theory have some connection with the topological-space structure, we shall build our σ-algebra out of topologically significant sets. There will be a choice for how to do so, and we come to that point in a moment.

We shall follow as much as possible the pattern of the development of Lebesgue measure on an interval of $\mathbb{R}^1$ or on all of $\mathbb{R}^1$, as occurred in Chapter V, in order to construct measures on X. The thing that is missing for general X occurs right at the start: it is the kind of geometric information that goes into regarding the length of an interval as a quantity worthy of study. That is where an ingenious idea comes into play, that of studying linear functionals on the vector space $C_{\text{com}}(X)$ of continuous scalar-valued functions on X that vanish off a compact subset of X. As in earlier chapters, it will not be important whether the scalars for $C_{\text{com}}(X)$ are real or complex, and the reader may fix attention on either of these.

On an interval $[a, b]$, we thus consider the space $C([a, b])$ of scalar-valued continuous functions on the interval. The particular linear functional of interest is the Riemann integral $\ell(f) = \mathcal{R}\int_a^b f(x)\,dx$, the notation with the $\mathcal{R}$ being as in Section VI.4. This kind of integral is a fairly simple object analytically; it was

[1] The measure-theoretic foundations of probability theory are discussed in the companion volume *Advanced Real Analysis*.

quickly shown to make sense in Theorem 1.26. Our point of view will be that the Riemann integral encodes information about the lengths of all intervals.

Why might one consider linear functionals? In the subject of linear algebra, linear functionals play an important role. Two important ways of realizing subsets of Euclidean space are parametric form and implicit form. In the case of a vector subspace of $\mathbb{R}^n$, the idea of parametric form leads us to represent the subspace as all linear combinations of members of a spanning set. If we use implicit form instead, the subspace is realized as all vectors satisfying a set of homogeneous linear equations, thus as the kernel of some linear function. The most primitive case of the latter is that there is just one nontrivial equation. Then the linear function has range the scalars, and the linear function is a linear functional. When there are several equations, the subspace is in effect described as the intersection of the kernels of several linear functionals.

Thus linear functionals in linear algebra arise in describing vector subspaces, specifically in describing subspaces by limiting their size from the outside. In analysis we have occasionally needed this kind of control of a subspace in proving theorems by an approximation argument. Two nontrivial examples were the proofs in Chapter VI of differentiation of integrals and the proof in Chapter IX of the boundedness of the Hilbert transform. In each case we proved a theorem for "nice" functions, and we obtained some estimate for all functions of interest. To connect the one conclusion with the other, we needed to know that the subspace of "nice" functions is dense. Corollary 6.4 was a result of this kind, saying that $C_{\text{com}}(\mathbb{R}^N)$ is dense in $L^1(\mathbb{R}^N)$ and in $L^2(\mathbb{R}^N)$. The proof given for Corollary 6.4 was more like an argument using spanning sets, showing that we can pass from $C_{\text{com}}(\mathbb{R}^N)$ to simple functions and then recalling that simple functions are dense as a consequence of basic properties of the Lebesgue integral.

However, we can visualize another argument of this kind, one with continuous linear functionals. If one could prove, for any proper closed vector subspace of our total space of functions (L^1 or L^2 or something else), that there is a nonzero continuous linear functional on the total space vanishing on the closed subspace, then we could test whether a given vector subspace is dense by examining the effect of continuous linear functionals when restricted to the subspace. Historically this idea began to be applied in analysis in the early part of the twentieth century at about the same time that people began thinking frequently about spaces of functions and not just individual functions. The key general existence tool for such continuous linear functionals was the Hahn–Banach Theorem, which we shall take up in Chapter XII.

In any event, out of this confluence of ideas arose the idea of considering continuous linear functionals on $C_{\text{com}}(X)$ as capturing enough information about X to make measure theory possible. The continuity of a linear functional will actually be somewhat concealed in what we do for most of this chapter, and instead we impose on the linear functional the natural condition that it needs to satisfy in order to provide a notion of integration—that it be ≥ 0 on functions ≥ 0.

Let us be more precise about the definitions. Let X be a locally compact Hausdorff space, and let $C_{\text{com}}(X)$ be the vector space of scalar-valued functions on X that vanish outside some compact set. For a specific function f, the **support** of f is the closure of the set where f is not zero. The members of $C_{\text{com}}(X)$ are then the continuous scalar-valued functions on X having compact support. A linear functional ℓ on $C_{\text{com}}(X)$ is said to be **positive** if $\ell(f) \geq 0$ whenever $f \geq 0$. The Riesz Representation Theorem, to be stated formally in Section 2 with all details in place, will say that to any such ℓ corresponds a measure μ on a certain σ-algebra of "topologically significant" sets such that

$$\ell(f) = \int_X f\,d\mu \qquad \text{for all } f \in C_{\text{com}}(X).$$

The "topologically significant" sets have to include the sets necessary to make each f in $C_{\text{com}}(X)$ measurable. At first glance it might seem that the smallest σ-algebra containing the open sets is the right object. But in fact this σ-algebra is unnecessarily large. In an uncountable discrete space, we do not need to have every subset measurable in order to have all the functions of compact support be measurable. Accordingly we define the σ-algebra $\mathcal{B}(X)$ of **Borel sets** of X to be the smallest σ-algebra containing all compact subsets of X.

The plan of attack now follows the steps in the construction of Lebesgue measure. We take the compact subsets of X to be the analog of the bounded intervals in $\mathbb{R}^1$, and we thus define the "elementary sets" in X to be the sets in the smallest ring $\mathcal{K}(X)$ containing all the compact sets. In the case of $\mathbb{R}^1$, every set in the ring generated by the bounded intervals is a finite disjoint union of sets that are the difference of two bounded intervals. We shall prove for X in Section 2 that every member of $\mathcal{K}(X)$ is a finite disjoint union of sets that are the difference of two compact sets.

For $\mathbb{R}^1$, we defined the measure of the difference of two bounded intervals to be the difference of their lengths as soon as the second interval is contained in the first; this was no loss of generality because the intersection of two bounded intervals is a bounded interval. The measure of a finite disjoint union was defined as the sum of the measures. We showed that this was well defined, and then we had a finite-valued nonnegative additive set function on a ring of sets.

For X, we define the measure of a compact set K by the natural formula

$$\mu(K) = \inf_{\substack{f \in C_{\text{com}}(X), \\ 0 \leq f \leq I_K}} \ell(f),$$

where I_K as usual is the indicator function of K. The intersection of two compact sets is compact, and thus we can define the measure of $K_1 - K_2$ for K_1 and K_2

compact, to be $\mu(K_1) - \mu(K_1 \cap K_2)$. We define the measure of the disjoint union of such sets $K_1 - K_2$ to be the sum of the measures. We have to prove that this is well defined, and then we have a finite-valued nonnegative additive set function μ on the ring $\mathcal{K}(X)$.

The next step for $\mathbb{R}^1$ was to prove complete additivity on the ring generated by the bounded intervals. With X, the problem is the same; we are to prove complete additivity on the ring $\mathcal{K}(X)$. Suppose that this has been done. Since μ is everywhere finite-valued on $\mathcal{K}(X)$, we can apply the Extension Theorem (Theorem 5.5) to extend μ to the generated σ-ring. Either this σ-ring is already the generated σ-algebra $\mathcal{B}(X)$, or Proposition 5.37 supplies a canonical extension to a measure on the generated σ-algebra $\mathcal{B}(X)$. This completes the construction of the measure μ on $\mathcal{B}(X)$. It is then a fairly easy matter to see that $\ell(f)$ is recovered as the integral of f if f is in $C_{\text{com}}(X)$: In the case of $\mathbb{R}^1$, we carried out this step by first establishing the Fundamental Theorem of Calculus for the Lebesgue integral of a continuous function; the argument appears at the end of Section V.3. A more direct argument would have been possible, and that direct argument works for general X.

Thus the problem comes down to proving that the set function, as defined on the ring of sets, is actually completely additive on that ring. In the case of $\mathbb{R}^1$, that complete additivity was an easy consequence of "regularity" of Lebesgue measure on the ring generated by the bounded intervals; in other words, the measure of any set in the ring could be approximated from within by the measure of compact sets in the ring and from without by the measure of open sets in the ring. Exactly the same approach works for general X, but the regularity has to be established.

Quantitatively the construction of the measure comes down to defining $\mu(K)$ for K compact as above and then proving three identities:

(i) $\mu(K_1) + \mu(K_2) = \mu(K_1 \cup K_2) + \mu(K_1 \cap K_2)$ if K_1 and K_2 are compact,

(ii) $\sup\limits_{\substack{f \in C_{\text{com}}(X), \\ 0 \le f \le I_U}} \ell(f) = \mu(K) - \mu(K - U)$ if U is any open set contained in some compact set K,

(iii) $\sup\limits_{\substack{K \subseteq U, \\ K \text{ compact}}} \mu(K) = \sup\limits_{\substack{f \in C_{\text{com}}(X), \\ 0 \le f \le I_U}} \ell(f)$ if U is open and has compact closure.

Identity (i) and an elementary but lengthy computation in elementary set theory together allow us to prove that μ is well defined on the ring $\mathcal{K}(X)$ under the definitions above. Once μ has been so extended, the right side of (ii) is just $\mu(U)$ if U is open with compact closure. Thus (iii) says that $\mu(U)$ is the supremum of $\mu(K)$ over compact sets K contained in U, provided U is open and has compact closure. Since $\mu(U)$ is trivially the infimum of $\mu(V)$ for open sets V in $\mathcal{K}(X)$ containing U, this is the regularity conclusion for U. It is easy to see that the subclass of $\mathcal{K}(X)$ for which regularity holds is a ring and contains the compact sets, and hence regularity is established for $\mathcal{K}(X)$.

When the locally compact Hausdorff space X is a metric space, the three identities above are fairly easy to prove. When X is metric, any indicator function I_K for K compact is the pointwise decreasing limit of members of $C_{\text{com}}(X)$ that are ≥ 0. In fact, if $D(\,\cdot\,, K)$ is the distance to K, then the sequence $\{f_n\}$ with $f_n(x) = \max\{0,\ 1-nD(x, K)\}$ has the required properties. A little trick proves in this case that $\mu(K) = \lim_n \ell(f_n)$. To prove (i), we choose such sequences $\{f_n\}$ and $\{g_n\}$ for K_1 and K_2. If φ is a member of $C_{\text{com}}(X)$ that is identically 1 on the union of the supports of f_1 and g_1, then $f_n + g_n = \min\{f_n + g_n, \varphi\} + \big(\max\{f_n + g_n, \varphi\} - \varphi\big)$ decomposes $f_n + g_n$ into the sum of such sequences for $K_1 \cup K_2$ and $K_1 \cap K_2$, and identity (i) follows from linearity of ℓ and a passage to the limit. Identities (ii) and (iii) follow from equally simple arguments.

The difficulty for a general locally compact Hausdorff space X is that the indicator function of a compact set need not be a pointwise decreasing limit of a sequence of continuous functions. The technicalities introduced by this fact have the effect of making the proofs of (i), (ii), and (iii) be more complicated, but these complications need not obscure the line of argument that is so clear in the metric case.

2. Riesz Representation Theorem

Throughout this section we fix the locally compact Hausdorff space X. We continue to let $C_{\text{com}}(X)$ be the space of continuous functions of compact support, $\mathcal{K}(X)$ be the ring of elementary sets, and $\mathcal{B}(X)$ be the σ-algebra of Borel sets.

A subset E of X is said to be **bounded** if it is contained in a compact set, hence if E^{cl} is compact; it is **σ-bounded** if it is contained in the countable union of compact sets. The class of all σ-bounded Borel sets is a σ-ring containing $\mathcal{K}(X)$, and it is therefore the smallest σ-ring containing $\mathcal{K}(X)$.

A measure on the Borel sets of X is called a **Borel measure** if it is finite on every compact set. A Borel measure μ is said to be **regular** if it satisfies

$$\mu(E) = \sup_{\substack{K \subseteq E,\\ K \text{ compact}}} \mu(K) \qquad \text{for every set } E \text{ in } \mathcal{B}(X)$$

$$\mu(E) = \inf_{\substack{U \supseteq E,\\ U \text{ open } \sigma\text{-bounded}}} \mu(U) \qquad \text{for every } \sigma\text{-bounded set } E \text{ in } \mathcal{B}(X).$$

Theorem 11.1 (Riesz Representation Theorem). If ℓ is a positive linear functional on $C_{\text{com}}(X)$, then there exists a unique regular Borel measure μ on X such that

$$\ell(f) = \int_X f\,d\mu \qquad \text{for all } f \in C_{\text{com}}(X).$$

EXAMPLES.

(1) If X is the line $\mathbb{R}^1$ and ℓ is given by Riemann integration $l(f) = \mathcal{R}\int_a^b f(x)\,dx$ whenever $[a, b]$ contains the support of f, then ℓ is a positive linear functional on $C_{\text{com}}(\mathbb{R}^1)$ and the corresponding μ is Lebesgue measure.

(2) If $X = S^2$ is the unit sphere in $\mathbb{R}^3$, parametrized by latitude θ_1 from 0 to π and by longitude θ_2 from 0 to 2π, then $\ell(f) = \mathcal{R}\int_0^\pi \int_0^{2\pi} f(\theta_1, \theta_2)\sin\theta_1\, d\theta_2\, d\theta_1$ is a positive linear functional on $C(S^2)$, and the corresponding measure, which is written $d\omega$ in the same way that Lebesgue measure is written as dx, is a rotation-invariant measure on the sphere such that $\int_{\mathbb{R}^3} F(x)\,dx = \int_0^\infty \int_{S^2} F(r\omega)r^2\,d\omega\,dr$ for every nonnegative Borel function on $\mathbb{R}^N$. The proof of this identity and of the rotation invariance will be indicated in Problem 5 at the end of the chapter.

(3) If X is general and if μ is a regular Borel measure on X, then $\ell(f) = \int_X f\,d\mu$ is a positive linear functional on $C_{\text{com}}(X)$.

The proof of Theorem 11.1 will occupy the remainder of this section. We begin with some lemmas clarifying the nature of the ring $\mathcal{K}(X)$, the linear functional ℓ, and general compact and open subsets of X. Then we recall the definition of $\mu(K)$ for compact sets and establish the identities (i), (ii), and (iii) in Section 1. Finally we give the details of how the three identities imply the theorem.

We begin with information about the ring $\mathcal{K}(X)$.

Lemma 11.2. The members of the ring $\mathcal{K}(X)$ are exactly all finite disjoint unions of subsets of V of the form $K - L$ with K and L compact and $L \subseteq K \subseteq V$. The ring $\mathcal{K}(X)$ may be characterized also as the smallest ring containing all bounded open subsets of X.

PROOF. If $K_1 - L_1$ and $K_2 - L_2$ are two sets of the same kind as V in the statement of the lemma, then the identity

$$\begin{aligned}&(K_1 - L_1) \cup (K_2 - L_2)\\ &= ((K_1 \cup K_2) - (L_1 \cup L_2)) \cup ((K_2 \cap L_1) - (L_1 \cap L_2)) \cup ((K_1 \cap L_2) - (L_1 \cap L_2))\end{aligned}$$

shows that a union of two such sets is a disjoint union, and the identity

$$(K_1 - L_1) - (K_2 - L_2) = ((K_1 \cap L_2) - (L_1 \cap L_2)) \cup (K_1 - (L_1 \cup (K_1 \cap K_2)))$$

shows that the difference of two such sets is such a set. Therefore the collection of all such sets is a ring of subsets of X. This ring contains all compact sets because any compact set K is of the form $K - \varnothing$, and hence this ring equals $\mathcal{K}(X)$.

Any open bounded set U is the difference of the compact sets U^{cl} and $U^{\text{cl}} - U$, and hence it lies in $\mathcal{K}(X)$. In the reverse direction Corollary 10.23 shows that any compact set K is contained in the interior L^o of some compact set L. Thus K is the difference of the bounded open sets L^o and $L^o - K$, and $\mathcal{K}(X)$ is contained in the smallest ring containing all bounded open sets.

Next we observe some properties of the linear functional ℓ. It is to be understood throughout the section that ℓ is a positive linear functional on $C_{\text{com}}(X)$. The positivity implies that $\ell(f-g) \geq 0$ if $f-g \geq 0$; the linearity therefore gives $\ell(f) \geq \ell(g)$ for $f \geq g$. The linear functional has a kind of continuity property, according to the following lemma.

Lemma 11.3. Let K be a compact set, and let $\{f_n\}$ be a sequence in $C_{\text{com}}(X)$ converging uniformly to a member f of $C_{\text{com}}(X)$ in such a way that $\text{support}(f_n) \subseteq K$ for all n. Then $\lim_n \ell(f_n)$ exists and equals $\ell(f)$.

PROOF. Corollaries 10.23 and 10.44 show that there exists a function F in $C_{\text{com}}(X)$ such that F takes values in $[0, 1]$ and is 1 on K. Since $f_n - f \leq |f_n - f|$ and $-(f_n - f) \leq |f_n - f|$, we have

$$|\ell(f_n) - \ell(f)| = |\ell(f_n - f)| \leq \ell(|f_n - f|) \leq \ell(c_n F) = c_n \ell(F),$$

where $c_n = \|f_n - f\|_{\text{sup}}$. The assumed uniform convergence means that c_n tends to 0. Since $\ell(F)$ is some fixed constant, the asserted convergence of $\ell(f_n)$ follows.

Lemma 11.4 (Dini's Theorem). If $\{f_n\}$ is a sequence of functions in $C_{\text{com}}(X)$ decreasing pointwise to 0, then $\{f_n\}$ converges uniformly to 0.

PROOF. Because of the pointwise decrease to 0, all the functions f_n have support contained in the compact set $K = \text{support}(f_1)$. Let $\epsilon > 0$ be given, and let U_n be the open set where the continuous function f_n is $< \epsilon$. The pointwise decrease implies that the U_n are increasing with n, and the limit of 0 implies that each x in K is in some U_n. Thus the open sets U_n form an open cover of K. By compactness, there is a finite subcover. Since the sets U_n are increasing, some particular U_N covers K. Then $\|f_n\|_{\text{sup}} \leq \epsilon$ for $n \geq N$.

The final step of preparation is to observe some properties of compact and open sets. A bounded subset of X is said to be a G_δ if it is the countable intersection of bounded open sets. It is said to be an F_σ if it is the countable union of compact sets. We shall be especially interested in compact G_δ's and in open bounded F_σ's.

Lemma 11.5. Let f be a member of $C_{\text{com}}(X)$ with values in $[0, 1]$. If $r > 0$, then the set where f is $\geq r$ is a compact G_δ. If $r \geq 0$, then the set where f is $> r$ is a bounded open F_σ.

PROOF. The set where f is $\geq r$ is closed because of continuity, and this closed set is a subset of the compact support. Hence the set is compact. Similarly the set where f is $> r$ is open because of continuity, and this open set is a subset of the compact support. Hence the set is bounded.

When $r \geq 0$, the set where f is $> r$ is the union, for $n \geq 1$, of the sets where f is $\geq r + \frac{1}{n}$. For $r > 0$ when N is large enough so that $r - \frac{1}{N} > 0$, the set where f is $\geq r$ is the intersection, for $n \geq N$, of the sets where f is $> r - \frac{1}{n}$. The lemma follows.

Lemma 11.6.

(a) If K is a compact G_δ, then there exists a decreasing sequence of bounded open sets U_n such that $U_n \supseteq U_{n+1}^{\text{cl}}$ for all n and $\cap_{n=1}^\infty U_n = K$.

(b) If U is a bounded open F_σ, then there exists an increasing sequence of compact sets K_n such that $K_n \subseteq K_{n+1}^o$ for all n and $\cup_{n=1}^\infty K_n = U$.

PROOF. For (a), let $\{V_n\}$ be a sequence of bounded open sets with intersection K. This is possible since K is a G_δ. Without loss of generality we may assume that the V_n decrease with n. We define the sequence $\{U_n\}$ inductively on n. Put $U_1 = V_1$. If U_n has been constructed, use Corollary 10.22 to find an open set V_n' such that $K \subseteq V_n'$ and $V_n'^{\text{cl}} \subseteq U_n$, and then define $U_{n+1} = V_n' \cap V_{n+1}$. Then the sets U_n have the required properties.

For (b), let $\{L_n\}$ be a sequence of compact sets with union U. This is possible since U is an F_σ. Without loss of generality we may assume that the L_n increase with n. We define the sequence $\{K_n\}$ inductively on n. Put $K_1 = L_1$. If K_n has been constructed, use Corollary 10.23 to find an open set V_n' such that $U \supseteq V_n'^{\text{cl}}$ and $V_n' \supseteq K_n$. The compact set $L_n' = V_n'^{\text{cl}}$ has $(L_n')^o \supseteq V_n'$. If we define $K_{n+1} = L_n' \cup L_{n+1}$, then the sets K_n have the required properties.

Lemma 11.7.

(a) If K is a compact G_δ, then there exists a decreasing sequence of functions f_n in $C_{\text{com}}(X)$ with values in $[0, 1]$ such that each f_n is 1 on some neighborhood of K and $\lim f_n = I_K$ pointwise.

(b) If U is a bounded open F_σ, then there exists an increasing sequence of functions f_n in $C_{\text{com}}(X)$ with values in $[0, 1]$ such that each f_n has compact support contained in U and $\lim f_n = I_U$ pointwise.

PROOF. For (a), apply Lemma 11.6a to choose a sequence of bounded open sets U_n with intersection K such that $U_n \supseteq U_{n+1}^{\text{cl}}$ for all n. Using Corollary 10.44, let g_n be a member of $C_{\text{com}}(X)$ with values in $[0, 1]$ such that g_n is 1 on U_{n+1}^{cl} and is 0 off U_n, and put $f_n = \min\{g_1, \ldots, g_n\}$. Then the functions f_n have the required properties.

For (b), apply Lemma 11.6b to choose a sequence of compact sets K_n with union U such that $K_n \subseteq K_{n+1}^o$ for all n. Using Corollary 10.44, let g_n be a member of $C_{\text{com}}(X)$ with values in $[0, 1]$ such that g_n is 1 on K_n and is 0 off K_{n+1}^o, and put $f_n = \max\{g_1, \ldots, g_n\}$. Then the functions f_n have the required properties.

Now we begin the proofs of the three identities in Section 1. If K is compact, let
$$\mu(K) = \inf\ \ell(f),$$
the infimum being taken over all f in $C_{\text{com}}(X)$ such that $f \geq I_K$. Since $\ell(\min\{f, 1\}) \leq \ell(f)$, there is no harm in considering only those f's taking values in $[0, 1]$. It is immediate from this definition and the positivity of ℓ that μ is nonnegative and **monotone** in the sense that $K' \subseteq K$ implies $\mu(K') \leq \mu(K)$. The next lemma is the key to being able to prove the three identities in Section 1.

Lemma 11.8. If K is a compact subset of X, then the infimum of $\ell(f)$ over all f in $C_{\text{com}}(X)$ such that $f \geq I_K$ equals the infimum of $\ell(f)$ over all f in $C_{\text{com}}(X)$ with values in $[0, 1]$ such that $f \geq I_N$ for some neighborhood N of K depending on f.

REMARK. In particular, $\mu(K)$ can be computed by using only functions $f \geq I_K$ that are equal to 1 in some neighborhood of K.

PROOF. The problem is to show that the first infimum I_1 is not less than the second infimum I_2. Let $\epsilon > 0$ be given. Choose f in $C_{\text{com}}(X)$ with values in $[0, 1]$ such that $f \geq I_K$ and $\ell(f) \leq I_1 + \epsilon$, and let L be the set where f is ≥ 1. Lemma 11.5 shows that L is a compact G_δ, and Lemma 11.7a produces a decreasing sequence of functions f_n in $C_{\text{com}}(X)$ with values in $[0, 1]$ such that each f_n is 1 on some neighborhood of L and $\lim f_n = I_L$ pointwise. Then the sequence $\{\max\{f_n, f\}\}$ is pointwise decreasing with limit $\max\{I_L, f\} = f$, and hence $\{\max\{f_n, f\} - f\}$ is a pointwise decreasing sequence in $C_{\text{com}}(X)$ with limit 0. By Dini's Theorem (Lemma 11.4), the sequence $\{\max\{f_n, f\} - f\}$ converges uniformly to 0, and hence $\ell(\max\{f_n, f\})$ decreases to $\ell(f)$. For some sufficiently large n_0, we therefore have $\ell(\max\{f_{n_0}, f\}) \leq I_1 + 2\epsilon$. The function $\max\{f_{n_0}, f\}$ is one of the functions that figures into I_2, and thus $I_2 \leq I_1 + 2\epsilon$. Since ϵ is arbitrary, $I_2 \leq I_1$.

Lemma 11.8 puts us in a position to prove identity (i) in Section 1 and to deduce that μ extends in a well-defined fashion to a nonnegative additive set function on $\mathcal{K}(X)$. We make use of the formula $a + b = \min\{a, b\} + \max\{a, b\}$, from which it follows that $a = \min\{a, b\} + (\max\{a, b\} - b)$.

Lemma 11.9. If K_1 and K_2 are any two compact subsets of X, then
$$\mu(K_1) + \mu(K_2) = \mu(K_1 \cup K_2) + \mu(K_1 \cap K_2).$$

REMARK. The argument in Lemma 11.8 adapts to give a quick proof of the present lemma when X is a metric space. In the metric case we can find a decreasing sequence $\{f_n\}$ of functions ≤ 1 in $C_{\text{com}}(X)$ with pointwise limit I_{K_1}. If

$f \geq I_{K_1}$, then the proof of Lemma 11.8 shows that $f_n f$ converges uniformly to f and hence $\ell(f_n f)$ decreases to $\ell(f)$. It follows that $\ell(f_n)$ decreases to $\mu(K_1)$ whenever f_n decreases to I_{K_1}. If we similarly choose $\{g_n\}$ decreasing to I_{K_2} and choose, by Corollary 10.44, a function $\varphi \in C_{\text{com}}(X)$ with values in [0, 1] that is identically 1 on the support of $f_1 + g_1$, then the formula stated just above shows that $f_n + g_n = \min\{f_n + g_n, \varphi\} + \big(\max\{f_n + g_n, \varphi\} - \varphi\big)$. The first term on the right side decreases pointwise to $I_{K_1 \cup K_2}$, and the second term decreases to $I_{K_1 \cap K_2}$. Thus a passage to the limit in the formula $\ell(f_n) + \ell(g_n) = \ell(\min\{f_n + g_n, \varphi\}) + \ell\big(\big(\max\{f_n + g_n, \varphi\} - \varphi\big)\big)$ immediately yields the result of the present lemma.

PROOF. Let f and g be functions in $C_{\text{com}}(X)$ with values in [0, 1] such that $f \geq I_{K_1}$ and $g \geq I_{K_2}$, and choose, by Corollary 10.44, $\varphi \in C_{\text{com}}(X)$ with values in [0, 1] that is identically 1 on the support of $f + g$. Then we have $f + g = \min\{f + g, \varphi\} + (\max\{f + g, \varphi\} - \varphi)$. The first term on the right side is $\geq I_{K_1 \cup K_2}$, and the second term is $\geq I_{K_1 \cap K_2}$. Therefore

$$\begin{aligned} \ell(f) + \ell(g) &= \ell(\min\{f + g, \varphi\}) + \ell((\max\{f + g, \varphi\} - \varphi)) \\ &\geq \mu(K_1 \cup K_2) + \mu(K_1 \cap K_2). \end{aligned}$$

Taking the infimum over f and then over g, we obtain

$$\mu(K_1) + \mu(K_2) \geq \mu(K_1 \cup K_2) + \mu(K_1 \cap K_2).$$

For the reverse direction let F be a member of $C_{\text{com}}(X)$ with values in [0, 1] that is $\geq I_{K_1 \cup K_2}$ and is equal to 1 at least on some open set U containing $K_1 \cup K_2$. Similarly let G be a member of $C_{\text{com}}(X)$ with values in [0, 1] that is $\geq I_{K_1 \cap K_2}$ and is equal to 1 at least on some open set V containing $K_1 \cap K_2$. Lemma 11.8 shows that F and G are the most general functions of a kind needed for the computation of $\mu(K_1 \cup K_2)$ and $\mu(K_1 \cap K_2)$. The sets U and V have compact closure in X since they are subsets of the supports of F and G. Choose, by Corollary 10.44, $\varphi \in C_{\text{com}}(X)$ with values in [0, 1] that is identically 1 on the support of $F + G$. Let V_0 be an open set with $K_1 \cap K_2 \subseteq V_0 \subseteq V_0^{\text{cl}} \subseteq V$. Then $(K_2 - V_0) \cap K_1 = K_2 \cap V_0^c \cap K_1 \subseteq V_0 \cap V_0^c = \varnothing$. So there exists an open set W such that $K_2 - V_0 \subseteq W \subseteq W^{\text{cl}} \subseteq K_1^c$.

We define f and g to be members of $C_{\text{com}}(X)$ having compact support contained in U and having with values in [0, 1] such that

$$f = \begin{cases} 1 & \text{on } K_1, \\ 0 & \text{on } W^{\text{cl}}, \end{cases}$$

and

$$g = \begin{cases} 1 & \text{on } K_2, \\ 0 & \text{on support}(f) - V. \end{cases}$$

The functions f and g exist by Corollary 10.44 if it is shown that the closed sets K_1 and W^{cl} are disjoint and the closed sets K_2 and $\text{support}(f) - V$ are disjoint. The sets K_1 and W^{cl} are disjoint since $W^{\text{cl}} \subseteq K_1^c$. For K_2 and $\text{support}(f) - V$, we observe that $\text{support}(f) \subseteq ((W^{\text{cl}})^c)^{\text{cl}} \subseteq (W^c)^{\text{cl}} = W^c \subseteq (K_2 - V_0)^c = V_0 \cup K_2^c \subseteq V \cup K_2^c$. Therefore

$$\begin{aligned}(\text{support}(f) - V) \cap K_2 &\subseteq (V \cup K_2^c) \cap V^c \cap K_2 \\ &= (V \cap V^c \cap K_2) \cup (K_2^c \cap V^c \cap K_2) = \varnothing.\end{aligned}$$

We conclude that f and g exist.

By inspection, $f \geq I_{K_1}$ and $g \geq I_{K_2}$, from which $f + g \geq I_{K_1} + I_{K_2}$. Then $\min\{f + g, \varphi\}$ is 1 on $K_1 \cup K_2$ and is 0 off U. Since F is 1 on U, we obtain

$$\min\{f + g, \varphi\} \leq F. \tag{$*$}$$

Since $f + g \geq I_{K_1} + I_{K_2} = I_{K_1 \cup K_2} + I_{K_1 \cap K_2}$, the function $\max\{f + g, \varphi\} - \varphi$ equals $f + g - 1$ on $K_1 \cup K_2$, and this in turn is ≤ 1 everywhere. Let us see that

$$\max\{f + g, \varphi\} - \varphi \leq G \tag{$**$}$$

everywhere. The only points x at which $(**)$ could possibly fail are those where $G(x) < 1$, hence points of V^c. At such points the definition of g shows that $f(x) + g(x) \leq 1$. If also x is in U, then $\varphi(x) = 1$ and we compute that $\max\{f(x) + g(x), \varphi(x)\} - \varphi(x) = 1 - 1 = 0$. Thus $(**)$ holds at points of $U \cap V^c$. At points of $U^c \cap V^c$, the equality $f(x) = g(x) = 0$ implies that $\max\{f(x) + g(x), \varphi(x)\} - \varphi(x) = \varphi(x) - \varphi(x) = 0$. Thus again $(**)$ holds, and hence $(**)$ holds at every point of V^c, therefore everywhere.

Addition of $(*)$ and $(**)$ gives $f + g \leq F + G$ everywhere. Therefore

$$\ell(F) + \ell(G) = \ell(F + G) \geq \ell(f + g) = \ell(f) + \ell(g) \geq \mu(K_1) + \mu(K_2).$$

Taking the infimum over F and then over G gives $\mu(K_1 \cup K_2) + \mu(K_1 \cap K_2) \geq \mu(K_1) + \mu(K_2)$ and completes the proof of the lemma.

Lemma 11.9 yields by iteration a corresponding formula with the sum of n terms on each side. This extension of Lemma 11.9 is a computation in Boolean algebra involving no analysis at all—only the fact that the collection of compact sets is closed under finite unions and intersections. The details are carried out in the next lemma.

Lemma 11.10. If $K_1 \dots, K_n$ are compact subsets of X, then

$$\sum_{l=1}^{n} \mu(K_l) = \sum_{k=1}^{n} \mu\Big(\bigcup_{1 \le i_1 < \cdots < i_k \le n} \Big(\bigcap_{j=1}^{k} K_{i_j} \Big)\Big).$$

PROOF. The argument is by induction on n, the base case of the induction being the case $n = 2$ that was settled by Lemma 11.9. Thus let $n > 2$, and assume the identity for the case $n - 1$. The inductive hypothesis gives

$$\sum_{l=1}^{n} \mu(K_l) = \sum_{k=1}^{n-1} \mu\Big(\bigcup_{1 \le i_1 < \cdots < i_k < n} \Big(\bigcap_{j=1}^{k} K_{i_j} \Big)\Big) + \mu(K_n). \qquad (*)$$

We shall prove by induction on $r \ge 1$ that

$$\begin{aligned}\sum_{l=1}^{n} \mu(K_l) = \sum_{k=1}^{r-1} \mu\Big(\bigcup_{1 \le i_1 < \cdots < i_k \le n} \Big(\bigcap_{j=1}^{k} K_{i_j} \Big)\Big) \\ + \mu\Big(\bigcup_{1 \le i_1 < \cdots < i_r = n} \Big(\bigcap_{j=1}^{r} K_{i_j} \Big)\Big) + \sum_{k=r}^{n-1} \mu\Big(\bigcup_{1 \le i_1 < \cdots < i_k < n} \Big(\bigcap_{j=1}^{k} K_{i_j} \Big)\Big),\end{aligned}$$

the base case of this induction being $r = 1$, where this identity reduces to $(*)$. The proof for the case $r = n$ will complete the inductive step for the outer induction and thereby will complete the proof of the lemma. To pass from r to $r + 1$ in the inner induction, the question is whether

$$\begin{aligned}\mu\Big(\bigcup_{1 \le i_1 < \cdots < i_r = n} \Big(\bigcap_{j=1}^{r} K_{i_j} \Big)\Big) + \mu\Big(\bigcup_{1 \le i_1 < \cdots < i_r < n} \Big(\bigcap_{j=1}^{r} K_{i_j} \Big)\Big) \\ \stackrel{?}{=} \mu\Big(\bigcup_{1 \le i_1 < \cdots < i_r \le n} \Big(\bigcap_{j=1}^{r} K_{i_j} \Big)\Big) + \mu\Big(\bigcup_{1 \le i_1 < \cdots < i_{r+1} = n} \Big(\bigcap_{j=1}^{r+1} K_{i_j} \Big)\Big).\end{aligned}$$

The union of the two sets on the left here is the first set on the right side. In view of Lemma 11.9, this formula will follow if it is shown that the second set on the right side is the intersection of the two sets on the left. The intersection of the two sets on the left side is equal to

$$\bigcup_{\substack{1 \le i_1 < \cdots < i_r = n, \\ 1 \le i'_1 < \cdots < i'_r < n}} \Big(\Big(\bigcap_{j=1}^{r} K_{i_j} \Big) \cap \Big(\bigcap_{j=1}^{r} K_{i'_j} \Big)\Big). \qquad (**)$$

A term in the union in this expression is an intersection of at least $r+1$ of the sets $K_1, \dots, K_n$, the last of which is K_n, namely the ones corresponding to indices $i'_1, \dots, i'_r$ and $i_r = n$. Every intersection of exactly $r + 1$ of the sets $K_1, \dots, K_n$ occurs if the last one is K_n because we can take $i_1 = i'_1, \dots, i_{r-1} = i'_{r-1}$. Any intersection of more than $r + 1$ sets is contained in one with exactly $r + 1$ sets, and thus $(**)$ equals $\bigcup_{1 \le i_1 < \cdots < i_{r+1} = n} \Big(\bigcap_{j=1}^{r+1} K_{i_j} \Big)$, as asserted.

A further formality is the derivation from these results that μ extends in a well-defined fashion to a nonnegative additive set function on the ring $\mathcal{K}(X)$. Again no analysis is involved, only the one additional fact that the intersection of two sets of the form $K - L$ with K and L compact is again of this form, specifically that $(K - L) \cap (K' - L') = (K \cap K') - (L \cup L')$.

Lemma 11.11. The set function μ extends in a well-defined fashion to a nonnegative additive set function on $\mathcal{K}(X)$ under the definition

$$\mu\Big(\bigcup_{j=1}^{n} (K_j - L_j)\Big) = \sum_{j=1}^{n} \big(\mu(K_j) - \mu(L_j)\big)$$

whenever K_j and L_j are compact with $L_j \subseteq K_j$ for each j with $1 \le j \le n$ and the sets $K_1 - L_1, \dots, K_n - L_n$ are pairwise disjoint.

REMARKS. Lemma 11.2 assures us that every member of $\mathcal{K}(X)$ is of the form in this lemma. The subtlety of the lemma arises from the fact that the sets K_j need not be disjoint.

PROOF. First let us see that μ is well defined in the case $j = 1$, i.e., that $K' - L' = K - L$ with $L' \subseteq K'$ and $L \subseteq K$ implies $\mu(K') - \mu(L') = \mu(K) - \mu(L)$. We are to show that $\mu(K') + \mu(L) = \mu(K) + \mu(L')$, and Lemma 11.9 shows that it is enough to show that $K' \cup L = K \cup L'$ and $K' \cap L = K \cap L'$. Suppose x is in $K' \cup L$. If x is in L, then x is in K, hence in $K \cup L'$. If x is in K' instead, then either x has to be in L' in the case that x is not in $K' - L'$ or x has to be in K in the case that x is in $K' - L' = K - L$. So $K' \cup L \subseteq K \cup L'$. If x is in $K' \cap L$, then x is not in $K - L$ and must be in L' in order to avoid being in $K' - L'$. So x is in $L \cup L' \subseteq K \cup L'$. Reversing the roles of $K' - L'$ and $K - L$, we see that $K' \cup L = K \cup L'$ and $K' \cap L = K \cap L'$.

Next suppose that $K' - L' = \bigcup_{j=1}^{n} (K_j - L_j)$ with $L' \subseteq K'$, $L_j \subseteq K_j$ for each j, and the sets $K_j - L_j$ disjoint. We are to show that $\mu(K') - \mu(L') = \sum_{j=1}^{n} (\mu(K_j) - \mu(L_j))$, i.e., that $\mu(K') + \sum_{j=1}^{n} \mu(L_j) = \mu(L') + \sum_{j=1}^{n} \mu(K_j)$. The argument will generalize that in the previous paragraph: The set $K' - L'$ has complement $L' \cup K'^c$, and therefore the given condition of disjointness means that

$$X = (L' \cup K'^c) \cup \bigcup_{j=1}^{n} (K_j - L_j) \qquad (*)$$

disjointly. Put $L_{n+1} = K'$ and $K_{n+1} = L'$, so that we are asking whether

$$\sum_{j=1}^{n+1} \mu(L_j) \stackrel{?}{=} \sum_{j=1}^{n+1} \mu(K_j).$$

In view of Lemma 11.10, it would be enough to show that

$$\bigcup_{1\le i_1<\cdots<i_{k+1}\le n+1}\Big(\bigcap_{j=1}^{k} L_{i_j}\Big) = \bigcup_{1\le i_1<\cdots<i_{k+1}\le n+1}\Big(\bigcap_{j=1}^{k} K_{i_j}\Big)$$

for $1 \le k \le n+1$. The left side is the set of x lying in at least k of the sets L_i, and the right side is the corresponding set for the K_i's. Thus it is enough to prove that the set of x lying in exactly r sets K_i is contained in the set of x lying in exactly r sets L_i, for $1 \le r \le n+1$.

We check this condition separately for the three cases $x \in L'$, $x \notin K'$, and $x \in K' - L'$. From $(*)$ we see that x in $L' \cup K'^c$ implies that x is not in any $K_j - L_j$ for $1 \le j \le n$. Hence for the first two cases, x is in L_j with $1 \le j \le n$ if and only if x is in K_j.

Case 1. $x \in L'$. For x to be in r of the sets $K_1, \dots, K_{n+1}$, x must be in $r-1$ of the sets $K_1, \dots, K_n$, hence in $r-1$ of the sets $L_1, \dots, L_n$. Since x is in L', it is in $K' = L_{n+1}$. Therefore x is in r of the sets $L_1, \dots, L_{n+1}$.

Case 2. $x \notin K'$. For x to be in r of the sets $K_1, \dots, K_{n+1}$, x must be in r of the sets $K_1, \dots, K_n$, hence in r of the sets $L_1, \dots, L_n$. Since x is not in K', it is not in L_{n+1}. Therefore x is in r of the sets $L_1, \dots, L_{n+1}$.

Case 3. $x \in K' - L'$. Since x is not in $L' \cup K'^c$, $(*)$ shows that x is in exactly one $K_j - L_j$ with $1 \le j \le n$. For x to be in r of the sets $K_1, \dots, K_{n+1}$, x must be in r of the sets $K_1, \dots, K_n$, hence in $r-1$ of the sets $L_1, \dots, L_n$. Since x is in $K' = L_{n+1}$, it is in r of the sets $L_1, \dots, L_{n+1}$.

For the general case, suppose that $\bigcup_{j=1}^m (K'_j - L'_j) = \bigcup_{j=1}^n (K_j - L_j)$. Intersecting both sides with $K'_i - L'_i$, we obtain

$$K'_i - L'_i = \bigcup_{j=1}^{n} ((K_j \cap K'_i) - ((L_j \cup L'_i) \cap (K_j \cap K'_i))).$$

The case just proved shows that

$$\mu(K'_i - L'_i) = \sum_{j=1}^{n} \big(\mu(K_j \cap K'_i) - \mu((L_j \cup L'_i) \cap (K_j \cap K'_i))\big)$$

and hence

$$\sum_{i=1}^{m} \mu(K'_i - L'_i) = \sum_{i=1}^{m}\sum_{j=1}^{n} \big(\mu(K_j \cap K'_i) - \mu((L_j \cup L'_i) \cap (K_j \cap K'_i))\big).$$

Similarly

$$\sum_{j=1}^{n} \mu(K_j - L_j) = \sum_{j=1}^{n}\sum_{i=1}^{m} \big(\mu(K_j \cap K'_i) - \mu((L_j \cup L'_i) \cap (K_j \cap K'_i))\big).$$

Therefore $\sum_{i=1}^m \mu(K'_i - L'_i) = \sum_{j=1}^n \mu(K_j - L_j)$, and the proof is complete.

In short order, we can now prove identities (ii) and (iii). Lemma 11.12 will prove (iii), and Lemma 11.13 will prove (ii).

Lemma 11.12. If U is any bounded open subset of X, then

$$\sup_{\substack{g\in C_{\text{com}}(X),\\ 0\le g\le I_U,\\ \text{support } g\subseteq U}} \ell(g) = \sup_{\substack{K\subseteq U,\\ K \text{ compact}}} \mu(K) = \sup_{\substack{f\in C_{\text{com}}(X),\\ 0\le f\le I_U}} \ell(f).$$

PROOF. Let S_1, S_2, S_3 be the three suprema in question. We first check that $S_1 \le S_2 \le S_3$. If g contributes to S_1, then $g \le I_{\text{support } g} \le I_U$. If $h \in C_{\text{com}}(X)$ has $I_{\text{support } g} \le h$, then $g \le h$ and hence $\ell(g) \le \ell(h)$. Taking the infimum over all such h, we obtain $\ell(g) \le \mu(\text{support } g) \le S_2$. Taking the supremum over all g therefore gives $S_1 \le S_2$. Next if K is compact with $K \subseteq U$, Corollary 10.44 allows us to find $f \in C_{\text{com}}(X)$ with values in [0, 1] such that f is equal to 1 on K and equal to 0 on U^c. Then $I_K \le f \le I_U$. The definitions of $\mu(K)$ and S_3 yield $\mu(K) \le \ell(f) \le S_3$. Taking the supremum over all K therefore gives $S_2 \le S_3$.

To complete the proof, we show that $S_1 \ge S_3$. Let $\epsilon > 0$ be given. Choose f in $C_{\text{com}}(X)$ such that $0 \le f \le I_U$ and $\ell(f) \ge S_3 - \epsilon$, and let V be the set where f is > 0. Lemma 11.5 shows that V is a bounded open F_σ, and Lemma 11.7b produces an increasing sequence of functions f_n in $C_{\text{com}}(X)$ with values in [0, 1], each with support some compact subset of V, such that $\lim f_n = I_V$ pointwise. Then the sequence $\{\min\{f_n, f\}\}$ is pointwise increasing with limit $\min\{I_V, f\}$. If x is a point where $I_V(x) < f(x)$, then $f(x) > 0$, x is in V, and $I_V(x) = 1$, contradiction. So there is no such point, and $\min\{I_V, f\} = f$. Therefore the sequence $\{f - \min\{f_n, f\}\}$ is a pointwise decreasing sequence in $C_{\text{com}}(X)$ with limit 0. By Dini's Theorem (Lemma 11.4), the sequence $\{f - \min\{f_n, f\}\}$ converges uniformly to 0, and hence $\ell(\min\{f_n, f\})$ increases to $\ell(f)$. For some sufficiently large n_0, we therefore have $\ell(\min\{f_{n_0}, f\}) \ge S_3 - 2\epsilon$. The function $\min\{f_{n_0}, f\}$ is one of the functions that figures into S_1, and thus $S_1 \ge \ell(\min\{f_{n_0}, f\}) \ge S_3 - 2\epsilon$. Since ϵ is arbitrary, $S_1 \ge S_3$.

Lemma 11.13. Let μ be extended to a nonnegative additive set function on $\mathcal{K}(X)$ as in Lemma 11.11. If U is a bounded open subset of X, then $\mu(U) = \sup_{K\subseteq U,\, K \text{ compact}} \mu(K)$.

PROOF. For the bounded open set U, let S_1, S_2, S_3 be the three equal suprema of Lemma 11.12. By definition, $\mu(U) = \mu(L) - \mu(L - U)$ for any compact set L containing U, and we are to prove that $\mu(U) = S_2$. If K is a compact subset of U, then $K \cup (L - U)$ is a disjoint union contained in L, and we have $\mu(K) + \mu(L - U) = \mu(K \cup (L - U)) \le \mu(L)$. Taking the supremum over all such K, we obtain $S_2 + \mu(L - U) \le \mu(L)$, i.e., $S_2 \le \mu(U)$.

Let h be any member of $C_{\text{com}}(X)$ with values in $[0, 1]$ such that $h \geq I_{L-U}$ and such that h is 1 on an open neighborhood N of $L - U$. Then $L \subseteq N \cup U$. For each point x of U, find an open neighborhood U_x of x with $U_x^{\text{cl}} \subseteq U$. Then N and the U_x's form an open cover of L, and there is a finite subcover. Let us say that $L \subseteq N \cup U_{x_1} \cup \cdots \cup U_{x_n}$. The set $K = U_{x_1}^{\text{cl}} \cup \cdots \cup U_{x_n}^{\text{cl}}$ is a compact subset of U, and $L \subseteq N \cup K$. Choose, by Corollary 10.44, a function $f \in C_{\text{com}}(X)$ with values in $[0, 1]$ such that f is 1 on K and is 0 off U. This function has $0 \leq f \leq I_U$. Since f is 1 on K and h is 1 on N, $h + f$ is ≥ 1 on L. Hence $\mu(L) \leq \ell(h + f) = \ell(h) + \ell(f) \leq \ell(h) + S_3$. Thus $\mu(L) \leq \mu(L - U) + S_3$ and $\mu(U) \leq S_3$. Since $S_3 = S_2$ by Lemma 11.12, $\mu(U) = S_2$ as required.

PROOF OF EXISTENCE IN THEOREM 11.1. If K is compact, we define $\mu(K)$, just as we did earlier in this section, to be the infimum of $\ell(f)$ over all f in $C_{\text{com}}(X)$ such that $f \geq I_K$. Lemma 11.11 shows that μ extends, necessarily in a unique fashion, to a well-defined nonnegative additive set function on $\mathcal{K}(X)$.

Consider the set $\mathcal{C}$ of all members E of $\mathcal{K}(X)$ satisfying the following regularity property: for each $\epsilon > 0$, there exist compact K and open bounded U with $K \subseteq E \subseteq U$ and $\mu(U - K) < \epsilon$. Lemma 11.13 shows that every open bounded set is in $\mathcal{C}$. We show closure of $\mathcal{C}$ under finite unions. If E_1 and E_2 are in $\mathcal{C}$, then we can choose K_1 and K_2 compact and U_1 and U_2 bounded open such that $K_1 \subseteq E_1 \subseteq U_1$, $K_2 \subseteq E_2 \subseteq U_2$, $\mu(U_1 - K_1) < \epsilon/2$, and $\mu(U_2 - K_2) < \epsilon/2$. Then $K_1 \cup K_2 \subseteq E_1 \cup E_2 \subseteq U_1 \cup U_2$ and $(U_1 \cup U_2) - (K_1 \cup K_2) \subseteq (U_1 - K_1) \cup (U_2 - K_2)$. It follows that $\mu((U_1 \cup U_2) - (K_1 \cup K_2)) \leq \mu((U_1 - K_1)) + \mu((U_2 - K_2)) < \epsilon$, and $\mathcal{C}$ is closed under finite unions.

We show closure of $\mathcal{C}$ under differences. If E_1 and E_2 are in $\mathcal{C}$, then we again choose K_1 and K_2 compact and U_1 and U_2 bounded open such that $K_1 \subseteq E_1 \subseteq U_1$, $K_2 \subseteq E_2 \subseteq U_2$, $\mu(U_1 - K_1) < \epsilon/2$, and $\mu(U_2 - K_2) < \epsilon/2$. Then $K_1 - U_2 \subseteq E_1 - E_2 \subseteq U_1 - K_2$, and $(U_1 - K_2) - (K_1 - U_2) \subseteq (U_1 - K_1) \cup (U_2 - K_2)$. Hence $\mu((U_1 - K_2) - (K_1 - U_2)) \leq \mu(U_1 - K_1) + \mu(U_2 - K_2) < \epsilon$, and $\mathcal{C}$ is closed under differences. By Lemma 11.2, $\mathcal{C}$ equals $\mathcal{K}(X)$. Thus every set in $\mathcal{K}(X)$ satisfies the regularity property.

Next let us see that μ is completely additive on $\mathcal{C}$. Let E_n be a disjoint sequence of sets in $\mathcal{K}(X)$ with union E in $\mathcal{K}(X)$. For every N, we have $\sum_{n=1}^{N} \mu(E_n) = \mu(E_1 \cup \cdots \cup E_N) \leq \mu(E)$. Hence $\sum_{n=1}^{\infty} \mu(E_n) \leq \mu(E)$. For the reverse inequality, let $\epsilon > 0$ be given. Choose, by the regularity property, K compact and U_n open bounded with $K \subseteq E$, $E_n \subseteq U_n$, $\mu(E - K) < \epsilon$, and $\mu(U_n - E_n) < \epsilon/2^n$. Then $K \subseteq E = \bigcup_{n=1}^{\infty} E_n \subseteq \bigcup_{n=1}^{\infty} U_n$. In other words, the sets U_n form an open cover of the compact set K. Some finite subcollection is a cover, and thus $K \subseteq U_1 \cup \cdots \cup U_N$ for some N. Then we have

$$\begin{aligned} \mu(E) &= \mu(E - K) + \mu(K) \leq \epsilon + \mu(U_1 \cup \cdots \cup U_N) \\ &\leq \epsilon + \textstyle\sum_{n=1}^{N} \mu(U_n) \leq \epsilon + \sum_{n=1}^{N} (\mu(E_n) + \epsilon/2^n) \leq \sum_{n=1}^{\infty} \mu(E_n) + 2\epsilon. \end{aligned}$$

Since ϵ is arbitrary, $\mu(E) \le \sum_{n=1}^{\infty} \mu(E_n)$. Therefore $\mu(E) = \sum_{n=1}^{\infty} \mu(E_n)$, and μ is completely additive on $\mathcal{K}(X)$.

The Extension Theorem (Theorem 5.5) shows that μ extends uniquely to a measure on the smallest σ-ring containing $\mathcal{K}(X)$, i.e., the σ-ring of σ-bounded Borel sets. Proposition 5.37 shows further that μ extends canonically to a measure on the σ-algebra of all Borel sets under the definition

$$\mu(E) = \sup_{\substack{F \subseteq E,\ F \in \mathcal{B}(X),\\ F\ \sigma\text{-bounded}}} \mu(F).$$

This defines μ on $\mathcal{B}(X)$. We are left with showing that μ is regular and that $\ell(f) = \int_X f\,d\mu$ for every $f \in C_{\text{com}}(X)$.

In showing that $\ell(f) = \int_X f\,d\mu$ for every $f \in C_{\text{com}}(X)$, it is enough to handle an arbitrary $f \ge 0$. Fix $\epsilon > 0$, and fix an integer N such that $\|f\|_{\sup} < N\epsilon$. For $0 \le n \le N$, define $f_n = \min\{f, n\epsilon\}$. Each f_n is in $C_{\text{com}}(X)$, the function f_0 is 0, and the function f_N is f. For $0 \le n < N$, define $g_n = f_{n+1} - f_n$. We can recover f from the g_n's as $f = \sum_{n=0}^{N-1} g_n$. For $n \ge 1$, define $K_n = \{x \mid f(x) \ge n\epsilon\}$, and let $K_0 = \text{support}(f)$. All the sets K_n are compact, and they decrease in size with n. In this notation the formula for g_n is

$$g_n(x) = \begin{cases} 0 & \text{if } x \notin K_n, \\ f(x) - n\epsilon & \text{if } x \in K_n - K_{n+1}, \\ \epsilon & \text{if } x \in K_{n+1}. \end{cases}$$

Consequently
$$\epsilon I_{K_{n+1}} \le g_n \le \epsilon I_{K_n}. \tag{$*$}$$

Integration therefore gives

$$\epsilon\mu(K_{n+1}) \le \int_X g_n\,d\mu \le \epsilon\mu(K_n). \tag{$\dagger$}$$

The inequality given as $I_{K_{n+1}} \le \epsilon^{-1} g_n$ in $(*)$ implies that $\mu(K_{n+1}) \le \epsilon^{-1}\ell(g_n)$. The other inequality $\epsilon^{-1} g_n \le I_{K_n}$ in $(*)$ says that any $h \in C_{\text{com}}(X)$ with $I_{K_n} \le h$ has $\epsilon^{-1} g_n \le h$. Taking the infimum over h yields $\epsilon^{-1}\ell(g_n) \le \mu(K_n)$. Thus we have

$$\epsilon\mu(K_{n+1}) \le \ell(g_n) \le \epsilon\mu(K_n). \tag{$\dagger\dagger$}$$

Subtracting $(\dagger)$ and $(\dagger\dagger)$, we obtain

$$-\epsilon(\mu(K_n) - \mu(K_{n+1})) \le \int_X g_n\,d\mu - \ell(g_n) \le \epsilon(\mu(K_n) - \mu(K_{n+1})).$$

Since $f = \sum_{n=0}^{N-1} g_n$, summing from $n = 0$ to $n = N - 1$ gives

$$\left| \int_X f\,d\mu - \ell(f) \right| \le \epsilon \sum_{n=0}^{N-1} (\mu(K_n) - \mu(K_{n+1})) = \epsilon\mu(\text{support}(f)).$$

Since ϵ is arbitrary, $\left|\int_X f\,d\mu - \ell(f)\right| = 0$. Thus $\ell(f) = \int_x f\,d\mu$.

Fix a compact subset K_0 of X, form the σ-ring $\mathcal{B}(X) \cap K_0$, and let $\mathcal{A}(K_0)$ be the collection of members E of $\mathcal{B}(X) \cap K_0$ such that $\mu(E)$ is the supremum of $\mu(K)$ over all compact subsets K of E and $\mu(E)$ is the infimum of $\mu(U)$ over all bounded open sets in X that contain E; the open sets in question need not lie within K_0. Since the sets in $\mathcal{A}(K_0)$ all have finite measure, the regularity condition on E is that there exist, for each $\epsilon > 0$, K compact and U bounded open with $K \subseteq E \subseteq U$ and $\mu(U - K) < \epsilon$. The same arguments as at the beginning of the present proof show that $\mathcal{A}(K_0)$ is closed under finite unions and differences. To see closure under countable disjoint unions, let $\{E_n\}$ be a disjoint sequence in $\mathcal{A}(K_0)$ with union E, let ϵ be given, and choose K_n compact and U_n bounded open with $K_n \subseteq E_n \subseteq U_n$ and $\mu(U_n - K_n) < \epsilon/2^n$. Applying Corollary 10.23, let L be a compact subset of X with $K_0 \subseteq L^o$. The sets K_n are disjoint, and thus $\sum_{n=1}^\infty \mu(K_n)$ converges. Choose N such that $\sum_{n=N+1}^\infty \mu(K_n) < \epsilon$. Define $U = L^o \cap \bigcup_{n=1}^\infty U_n$, $K = \bigcup_{n=1}^N K_n$, $K_\infty = \bigcup_{n=1}^\infty K_n$, and $F = \bigcup_{n=N+1}^\infty K_n$. Then K is compact, U is bounded open, and $K \subseteq E \subseteq U$. Since $K_\infty = K \cup F$, we have

$$\begin{aligned}\mu(U - K) &\le \mu(U - K_\infty) + \mu(F) \le \mu\Big(\bigcup_{n=1}^\infty (U_n - K_n)\Big) + \mu\Big(\bigcup_{n=N+1}^\infty K_n\Big)\\ &\le \sum_{n=1}^\infty \mu(U_n - K_n) + \sum_{n=N+1}^\infty \mu(K_n) \le \sum_{n=1}^\infty \epsilon/2^n + \epsilon = 2\epsilon.\end{aligned}$$

Thus $\mathcal{A}(K_0)$ is closed under countable disjoint unions and is a σ-ring. Since the compact subsets of K_0 are in $\mathcal{A}(K_0)$, we conclude that $\mathcal{A}(K_0) = \mathcal{B}(K_0)$.

This proves regularity for all bounded sets. If E is σ-bounded, we can choose an increasing sequence $\{L_n\}$ of compact sets whose union contains E. Put $E_n = E \cap L_n$. Given $\epsilon > 0$, we apply the previous step to choose K_n compact and U_n bounded open such that $K_n \subseteq E_n \subseteq U_n$ and $\mu(U_n - K_n) < \epsilon/2^n$. Taking $U = \bigcup_{n=1}^\infty U_n$ and $K_\infty = \bigcup_{n=1}^\infty K_n$, we have $K_\infty \subseteq E \subseteq U$ and $\mu(U - K_\infty) < \epsilon$. Thus $\mu(U) \le \mu(E) + \epsilon$, and $\mu(E) \le \mu(K_\infty) + \epsilon$. The first of these inequalities, being possible for any ϵ, shows that $\mu(E)$ is the infimum of the measures of open σ-bounded sets containing E. Since $\mu(K_\infty) = \lim_N \mu\big(\bigcup_{n=1}^N K_n\big)$ by complete additivity, the second of these inequalities, being possible for any ϵ, shows that $\mu(E)$ is the supremum of the measures of compact sets contained in E.

This proves regularity for all σ-bounded sets. If E is a Borel set that is not σ-bounded, we know that $\mu(E)$ is the supremum of the measures of $\mu(F)$ for σ-bounded Borel subsets F of E, and we know that $\mu(F)$ is the supremum of the measures of $\mu(K)$ for compact subsets K of F. Therefore $\mu(E)$ is the supremum of the measures of $\mu(K)$ for compact subsets K of E. This completes the proof of regularity of μ.

PROOF OF UNIQUENESS IN THEOREM 11.1. Let μ be the constructed measure, and let ν be a second measure satisfying the properties of the theorem. The assumed regularity of ν implies that it is enough to prove that $\nu(K) = \mu(K)$ for every compact subset K of X. Fix K, and let α be the infimum defining $\mu(K)$, namely the infimum of $\ell(f)$ over all $f \in C_{\text{com}}(X)$ with values in [0, 1] such that $I_K \leq f$. Integrating this inequality with respect to ν, we see that $\nu(K) \leq \int_X f\,d\nu$ and therefore $\nu(K) \leq \alpha$. Suppose that $\nu(K) < \alpha$. By Corollary 10.23 and the assumed regularity of ν, we can find a bounded open set U with $U \supseteq K$ and $\nu(U) < \alpha$. By Corollary 10.44 we can find a function $g \in C_{\text{com}}(X)$ with values in [0, 1] such that g is 1 on K and is 0 off U. Then $I_K \leq g \leq I_U$. Hence $\ell(g) = \int_X g\,d\mu = \int_X g\,d\nu \leq \int_X I_U\,d\nu = \nu(U) < \alpha \leq \ell(g)$, and we obtain a contradiction. We conclude that $\nu(K) = \alpha = \mu(K)$, and the uniqueness follows.

3. Regular Borel Measures

The fact that compact sets for a general locally compact Hausdorff X need not be countable intersections of open sets suggests a look at the ring of sets generated by the compact sets that are indeed such intersections, as well as the associated σ-algebra. The sets in this σ-algebra are known as "Baire sets," and it turns out that the members of $C_{\text{com}}(X)$ are measurable with respect to this σ-algebra. The σ-algebra of Baire sets can be strictly smaller than the σ-algebra of Borel sets, and thus one can make a case for limiting oneself to Baire sets all along. This would be a fine point, one not worth pursuing here, but for one fact: the σ-algebra of Baire sets for $X \times Y$ is a correct σ-algebra to use in Fubini's Theorem for changing iterated integrals over X and Y to a double integral—and this may not be true when Borel sets are used.

This fact about Fubini's Theorem might seem to be a telling argument for replacing Borel sets by Baire sets everywhere in the theory. The difficulty is that it is a little tedious to check constantly whether sets are Baire sets—for example, whether one-point sets are Baire sets. Thus the normal practice is to work with Borel sets and to resort to Baire sets only when Fubini's Theorem comes into play in a way that makes the distinction important. The most frequent case that arises in applications of Fubini's Theorem in this theory is that a function on $X \times Y$ is continuous with compact support, in which case only Baire sets are involved anyway.

Thus let X be a locally compact Hausdorff space. The sets in the smallest σ-algebra $\mathcal{B}(X)$ containing the compact sets are the **Borel sets**, and the sets in the smallest σ-algebra $\mathcal{B}_0(X)$ containing the compact G_δ's are the **Baire sets**. Measurable functions in the first case will be called **Borel measurable functions** or **Borel functions**, and measurable functions in the second case will be called

Baire measurable functions or **Baire functions**. We shall observe in Corollary 11.16 below that every member of $C_{\text{com}}(X)$ is a Baire function.

If the locally compact Hausdorff space X is a metric space, then any closed set F is the intersection of the sets $U_n = \{x \mid D(x, F) < \frac{1}{n}\}$, where $D(\,\cdot\,, F)$ is the distance to the set F. Consequently every compact subset of X is a G_δ, and every Borel set is a Baire set.

Proposition 11.14. If K and U are subsets of X with K compact, U open, and $K \subseteq U$, then there exist a compact G_δ, say K_0, and an open bounded F_σ, say U_0, such that $K \subseteq U_0 \subseteq K_0 \subseteq U$.

PROOF. Choose by Corollary 10.44 a member f of $C_{\text{com}}(X)$ with values in $[0, 1]$ such that f is 1 on K and is 0 on U^c. If K_0 is the set where f is $\geq \frac{1}{2}$ and U_0 is the set where f is $> \frac{1}{2}$, then Lemma 11.5 shows that K_0 and U_0 have the required properties.

Corollary 11.15. Any σ-compact open subset of X is a Baire set.

PROOF. If $U = \bigcup_{n=1}^\infty K_n$ is open with each K_n compact, we can apply Proposition 11.14 to the inclusion $K_n \subseteq U$ and find a set $(K_n)_0$ that is a compact G_δ and has $K_n \subseteq (K_n)_0 \subseteq U$. Then $U = \bigcup_{n=1}^\infty (K_n)_0$ exhibits U as the countable union of compact G_δ's, hence as a Baire set.

Corollary 11.16. Every member of $C_{\text{com}}(X)$ is a Baire function.

PROOF. This is immediate from Lemma 11.5 and Corollary 11.15.

Proposition 11.17. If X and Y are σ-compact, then the product σ-algebra for $X \times Y$ obtained from the Baire sets of X and Y is the σ-algebra of Baire sets of $X \times Y$.

PROOF. If K_X and K_Y are compact G_δ's in X and Y, then $K_X \times K_Y$ is a compact G_δ in $X \times Y$, and it follows that $\mathcal{B}_0(X) \times \mathcal{B}_0(Y) \subseteq \mathcal{B}_0(X \times Y)$. For the reverse inclusion let K be a compact G_δ in $X \times Y$, and write K as $K = \bigcap_{n=1}^\infty U_n$ with each U_n open. We construct open sets S_n in $\mathcal{B}_0(X) \times \mathcal{B}_0(Y)$ with $K \subseteq S_n \subseteq U_n$, and then it follows that $K = \bigcap_{n=1}^\infty S_n$ and K is a Baire set.

To do so, it is enough to show that if $K \subseteq W$ with W open, then there is an open set S in $\mathcal{B}_0(X) \times \mathcal{B}_0(Y)$ with $K \subseteq S \subseteq W$. For each (x, y) in K, find open neighborhoods U_x of x and V_y of y such that $U_x \times V_y \subseteq W$. Proposition 11.14, applied to the inclusion $\{x\} \subseteq U_x$ and then to the inclusion $\{y\} \subseteq V_y$, shows that we may assume that U_x and V_y are open F_σ's. In view of Corollary 11.15, they are then Baire sets. Hence $U_x \times V_y$ is in $\mathcal{B}_0(X) \times \mathcal{B}_0(Y)$. As (x, y) varies, the sets $U_x \times V_y$ form an open cover of K, and there is a finite subcover. We can take S to be the union of the elements in the finite subcover, and then S has the required properties.

Now we turn our attention to measures. A **Baire measure** on X is a measure on the Baire sets that is finite on every compact G_δ. The restriction of a Borel measure to the Baire sets is a Baire measure. We are going to prove that Baire measures are automatically regular in the same sense that Borel measures in $\mathbb{R}^N$ are automatically regular.

Proposition 11.18. Every Baire measure μ is regular in the following sense:

$$\mu(E) = \sup_{\substack{K\subseteq E,\\ K \text{ compact } G_\delta}} \mu(K) \qquad \text{for every set } E \text{ in } \mathcal{B}_0(X),$$

$$\mu(E) = \inf_{\substack{U\supseteq E,\\ U \text{ open } F_\sigma}} \mu(U) \qquad \text{for every } \sigma\text{-bounded set } E \text{ in } \mathcal{B}_0(X).$$

REMARK. Since Baire sets and Borel sets are the same in a metric space, this proposition generalizes the known regularity of Borel measures on any open subset of $\mathbb{R}^n$, as given in Theorem 6.25.

PROOF. If L is a compact G_δ, then $\mu(L)$ is certainly the supremum of $\mu(K)$ for the compact G_δ's contained in L. Suppose that U is σ-bounded open with $L \subseteq U$. Proposition 11.14 produces a bounded open set U_0 that is an F_σ and has $L \subseteq U_0 \subseteq U$. Consequently $\mu(L)$ is the infimum of $\mu(U_0)$ for the open F_σ's containing L. Thus every compact G_δ satisfies the stated regularity condition.

The remainder of the proof runs parallel to the proof of regularity at the end of the proof of existence for Theorem 11.1, and we shall be brief. Fix a compact G_δ in X, say K_0. Form the σ-ring $\mathcal{B}_0(X) \cap K_0$, and let $\mathcal{A}_0(K_0)$ be the collection of members E of $\mathcal{B}_0(X) \cap K_0$ such that $\mu(E)$ is the supremum of $\mu(K)$ over all compact subsets K of E that are G_δ's and $\mu(E)$ is the infimum of $\mu(U)$ over all open supersets U of E that are F_σ's; the open sets in question need not lie within K_0. Since the sets in $\mathcal{A}_0(K_0)$ all have finite measure, the regularity condition on E is that there exist, for each $\epsilon > 0$, K compact and U open of the correct kind with $K \subseteq E \subseteq U$ and $\mu(U - K) < \epsilon$. The same arguments as earlier show that $\mathcal{A}_0(K_0)$ is closed first under finite unions and differences, then under countable disjoint unions. Thus $\mathcal{A}_0(K_0)$ is a σ-ring containing all compact G_δ's, and we conclude that $\mathcal{A}(K_0) = \mathcal{B}(K_0)$.

This proves regularity for all bounded Baire sets. If the Baire set E is σ-bounded, we can choose an increasing sequence $\{L_n\}$ of compact G_δ's whose union contains E. Put $E_n = E \cap L_n$. Then the same argument as earlier, using the sets E_n, shows that the regularity condition holds for E.

Finally if E is a Baire set that is not σ-bounded, we know that $\mu(E)$ is the supremum of the measures of $\mu(F)$ for σ-bounded Baire subsets F of E, and we know that $\mu(F)$ is the supremum of the measures of $\mu(K)$ for compact subsets K of F that are G_δ's. Therefore $\mu(E)$ is the supremum of the measures of $\mu(K)$ for compact subsets K of E that are G_δ's.

Proposition 11.19. If ν is a Baire measure on X, then there is one and only one regular Borel measure μ on X whose restriction to the Baire sets is μ.

PROOF. Since the members of $C_{\text{com}}(X)$ are Baire functions (Corollary 11.16), we can define a positive linear functional ℓ on $C_{\text{com}}(X)$ by $\ell(f) = \int_X f\,d\nu$. The uniqueness of the extending μ follows from the uniqueness part of Theorem 11.1. For existence we take μ to be the regular Borel measure given by the existence part of Theorem 11.1. We are to prove that μ and ν agree on Baire sets. The measures μ and ν agree on compact G_δ's by Lemma 11.7a and dominated convergence. By regularity of Baire measures (Proposition 11.18), μ and ν agree on all Baire sets.

Proposition 11.20. Suppose that X is compact and that μ and ν are Borel measures on X with μ regular. If ν is absolutely continuous with respect to μ, then ν is regular.

PROOF. Let $\epsilon > 0$ be given. The Radon–Nikodym Theorem (Theorem 9.16) and Corollary 5.24 together show that there exists $\delta > 0$ such that any Borel set A with $\mu(A) < \delta$ has $\nu(A) < \epsilon$. Let E be a Borel set to be tested for regularity under ν. Since μ is regular, we can choose K compact and U open with $K \subseteq E \subseteq U$ and $\mu(U - K) < \delta$. Then $\nu(U - K) < \epsilon$, and it follows that $\nu(E)$ is approximated within ϵ by $\nu(K)$ and $\nu(U)$.

Proposition 11.21. If μ is a regular Borel measure on X and if $1 \le p < \infty$, then

(a) $C_{\text{com}}(X)$ is dense in $L^p(X, \mu)$,
(b) the smallest closed subspace of $L^p(X, \mu)$ containing all indicator functions of compact G_δ's in X is $L^p(X, \mu)$ itself.

REMARK. This generalizes conclusions (a) and (b) of Proposition 9.9 from open subsets of $\mathbb{R}^N$ to all locally compact Hausdorff spaces.

PROOF. If E is a Borel set of finite μ measure and if ϵ is given, the regularity of μ allows us to choose a compact set K with $K \subseteq E$ and $\mu(E - K) < \epsilon$. Then we can find a bounded open set U with $K \subseteq U$ and $\mu(U - K) < \epsilon$, and Proposition 11.14 gives us a compact G_δ set K_0 such that $K \subseteq K_0 \subseteq U$. We have $\int_{\mathbb{R}^N} |I_E - I_K|^p\,d\mu = \mu(E - K) < \epsilon$, $\int_{\mathbb{R}^N} |I_U - I_K|^p\,d\mu = \mu(U - K) < \epsilon$, and $\int_{\mathbb{R}^N} |I_U - I_{K_0}|^p\,d\mu = \mu(U - K_0) < \epsilon$. Consequently we see in succession that the closure in $L^p(X, \mu)$ of the set of all indicator functions of compact sets contains all indicator functions of Borel sets of finite μ measure, the closure in $L^p(X, \mu)$ of the set of all indicator functions of bounded open sets contains all indicator functions of Borel sets of finite μ measure, and the closure in $L^p(X, \mu)$ of the set of all indicator functions of compact G_δ's contains all indicator functions of Borel sets of finite μ measure. Proposition 5.56 shows consequently that the smallest closed subspace of $L^p(X, \mu)$ containing all indicator functions of compact Baire sets is $L^p(X, \mu)$ itself. This proves (b).

For (a), let K_0 be a compact G_δ, and use Lemma 11.7a to choose a decreasing sequence $\{f_n\}$ of real-valued members of $C_{\text{com}}(\mathbb{R}^N)$ with pointwise limit I_{K_0}. Since f_1^p is integrable, dominated convergence yields $\lim_n \int_{\mathbb{R}^N} |f_n - I_{K_0}|^p \, d\mu = 0$. Hence the closure of $C_{\text{com}}(X)$ in $L^p(X, \mu)$ contains all indicator functions of compact G_δ's. By Proposition 5.55d this closure contains the smallest closed subspace of $L^p(X, \mu)$ containing all indicator functions of compact G_δ's. Conclusion (b) shows that the latter subspace is $L^p(X, \mu)$ itself. This proves (a).

Corollary 11.22. Suppose that X is a locally compact separable metric space. If μ is a Borel measure on X and if $1 \le p < \infty$, then

(a) $C_{\text{com}}(X)$, as a normed linear space under the supremum norm, is separable,

(b) $L^p(X, \mu)$ is separable.

REMARK. This generalizes Corollary 6.27c and Proposition 9.9c from open subsets of $\mathbb{R}^N$ to all locally compact separable metric spaces. The measure μ is automatically regular by Proposition 11.8 since Baire measures and Borel measures coincide in any locally compact metric space.

PROOF. Part (a) is proved by the same argument as for Corollary 6.27c. What is required is a substitute for Lemma 6.22a in order to obtain a sequence $\{F_n\}_{n=1}^\infty$ of compact subsets of X with union X such that $F_n \subseteq F_{n+1}^o$ for all n. It was observed at the beginning of Section X.3 that separable implies Lindelöf, and it follows from Proposition 10.24 that X is consequently σ-compact. Application of Proposition 10.25 then gives the sequence $\{F_n\}_{n=1}^\infty$. Corollary 2.59 is still to be applied to $C(F_n)$; since F_n is a compact metric space, the corollary shows that $C(F_n)$ is separable, and the argument goes through.

Part (b) follows from (a) and Proposition 11.21a in the same way that Corollary 6.27d follows from parts (a) and (c) of that corollary. The sequence $\{F_n\}_{n=1}^\infty$ of the previous paragraph is to be used in the argument.

Theorem 11.23 (Helly–Bray Theorem). Let X be a locally compact separable metric space. If $\{\mu_n\}$ is a sequence of Borel measures on X with $\{\mu_n(X)\}$ bounded, say by M, then there exist a Borel measure μ on X and a subsequence $\{\mu_{n_k}\}$ such that $\mu(X) \le M$ and $\lim_n \int_X f \, d\mu_{n_k} = \int_X f \, d\mu$ for all f in $C_{\text{com}}(X)$.

REMARKS. In the terminology of Section V.9, the measures μ_n are continuous linear functionals on the normed linear space $C_{\text{com}}(X)$, and the norm of the linear functional corresponding to μ_n is $\mu_n(X)$. The convergence is weak-star convergence, and the limiting linear functional is given by a Borel measure μ

with $\mu(X) \le M$. The theorem amounts to an application of the preliminary form of Alaoglu's Theorem (Theorem 5.58) and the identification of the limit as a measure.

PROOF. The proof consists of filling in the details in the remarks above. We regard $Y = C_{\text{com}}(X)$ as a normed linear space with the supremum norm. Any Borel measure ν on X defines by integration a linear functional on Y with norm given by $\|\nu\| = \sup_{f \in C_{\text{com}}(X),\ \|f\| \le 1} \left| \int_X f\,d\nu \right|$. The right side is certainly $\le \|f\|_{\sup} \nu(X)$. In the reverse direction, let $\{K_n\}$ be an increasing sequence of compact subsets of X with union X, so that $\lim_n \nu(K_n) = \nu(X)$. Choose functions $f_n : X \to [0, 1]$ in $C_{\text{com}}(X)$ by Corollary 10.44 such that f_n is 1 on K_n. Then $\|f_n\|_{\sup} \le 1$ for all n, and $\int_X f_n\,d\nu \ge \int_{K_n} d\nu = \nu(K_n)$. Hence $\|\nu\| \ge \limsup_n \nu(K_n) = \nu(X)$, and we conclude that $\|\nu\| = \nu(X)$.

Thus the given sequence $\{\mu_n\}$ corresponds to a sequence in Y^* with $\|\mu_n\| \le M$ for all n. Corollary 11.22 shows that Y is separable. Theorem 5.58 therefore applies and yields a subsequence $\{\mu_{n_k}\}$ and a member ℓ of Y^* with $\|\ell\| \le M$ such that $\lim_k \int_X f\,d\mu_{n_k} = \ell(f)$ for all f in $C_{\text{com}}(X)$. If $f \ge 0$, $\lim_k \int_X f\,d\mu_{n_k}$ is certainly ≥ 0, and thus ℓ is a positive linear functional on $C_{\text{com}}(X)$. The Riesz Representation Theorem (Theorem 11.1) produces a Borel measure μ on X with $\ell(f) = \int_X f\,d\mu$ for all f in $C_{\text{com}}(X)$. Since $\|\ell\| \le M$, we have $\mu(X) \le M$.

4. Dual to Space of Finite Signed Measures

We continue in this section with X as a locally compact Hausdorff space. We now change the point of view a little and regard $C_{\text{com}}(X)$ as a normed linear space under the supremum norm $\|f\|_{\sup} = \sup_{x \in X} |f(x)|$. The problem is to identify all continuous linear functionals on this normed linear space. We shall see shortly that it is enough to handle the case that X is compact.

If X^* is the one-point compactification of X, then two spaces to be considered in conjunction with $C_{\text{com}}(X)$ are $C(X^*)$, the space of continuous scalar-valued functions on X^*, and $C_0(X)$, the space of continuous scalar-valued functions on X that "vanish at infinity." When applied to a function f, the term **vanishes at infinity** means that for any $\epsilon > 0$, there is some compact set with the property that $|f(x)| \le \epsilon$ outside that set. It is equivalent to say that f extends to a member of $C(X^*)$ that is 0 at ∞.

The three spaces $C_{\text{com}}(X)$, $C_0(X)$, and $C(X^*)$ are related. In the first place, $C_{\text{com}}(X)$ is dense in $C_0(X)$. In fact, if f is in $C_0(X)$ and if $\epsilon > 0$ is given, we find K compact with $|f(x)| \le \epsilon$ outside K. Corollary 10.44 supplies a member g of $C_{\text{com}}(X)$ with values in [0, 1] that is 1 on K. Then the product fg is in $C_{\text{com}}(X)$, and $\|f - fg\|_{\sup} \le \epsilon$. Thus $C_{\text{com}}(X)$ is dense in $C_0(X)$. Any continuous linear functional on $C_{\text{com}}(X)$ is uniformly continuous by Proposition 5.57, and

Proposition 2.47 shows that it extends uniquely to a continuous linear functional on $C_0(X)$. Thus the continuous linear functionals on $C_0(X)$ and $C_{\text{com}}(X)$ are in one-one correspondence by restriction.

If we identify $C_0(X)$ as the subspace of $C(X^*)$ of functions equal to 0 at ∞, then every continuous linear functional on $C(X^*)$ restricts to a continuous linear functional on $C_0(X)$. In the reverse direction every continuous linear functional on $C_0(X)$ extends (nonuniquely) to a continuous linear functional on $C(X^*)$. In fact, let ℓ_0 be a continuous linear functional on $C_0(X)$, and fix a member f_0 of $C(X^*)$ with $f_0(\infty) = 1$. If f is any member of $C(X^*)$, then $f - f(\infty)f_0$ is in $C_0(X)$ and it makes sense to define $\ell(f) = \ell_0(f - f(\infty)f_0)$. Since

$$\begin{aligned}|\ell(f)| = |\ell_0(f - f(\infty)f_0)| &\le \|\ell_0\|\|f - f(\infty)f_0\|_{\text{sup}} \\ &\le \|\ell_0\|(\|f\|_{\text{sup}} + |f(\infty)|\|f_0\|_{\text{sup}}) \le \|\ell_0\|(1 + \|f_0\|_{\text{sup}})\|f\|_{\text{sup}},\end{aligned}$$

ℓ is bounded on $C(X^*)$ and is therefore continuous. Thus the study of continuous linear functionals on $C_{\text{com}}(X)$ reduces to the case that X is compact.

The first result below shows that any continuous linear functional on $C(X)$ with X compact is a finite linear combination of positive linear functionals. In view of Theorem 11.1, it is therefore given as a finite linear combination of integrations with respect to regular Borel measures. The remainder of the section will be devoted to making this result look tidier and seeing what happens to various norms under the correspondence.

Proposition 11.24. Let X be a compact Hausdorff space, and let ℓ be a continuous linear functional on $C(X)$. If ℓ takes real values on real-valued functions, define, for $f \ge 0$ in $C(X)$,

$$\ell^+(f) = \sup_{0 \le g \le f} \ell(g) \qquad \text{and} \qquad \ell^-(f) = \ell^+(f) - \ell(f);$$

then ℓ^+ and ℓ^- extend to positive linear functionals on $C(X)$ such that $\ell = \ell^+ - \ell^-$. If ℓ does not necessarily take real values on real-valued functions, then ℓ is a complex linear combination of positive linear functionals on $C(X)$.

PROOF. The functions f and g in this argument will all be in $C(X)$. For general ℓ not necessarily taking real values on real-valued functions, define $\bar\ell(f) = \overline{\ell(\bar f)}$. We readily check that $\bar\ell$ is a continuous linear functional on $C(X)$, that $\ell_R = \frac{1}{2}(\ell + \bar\ell)$ and $\ell_I = \frac{1}{2i}(\ell - \bar\ell)$ are continuous linear functionals on $C(X)$ taking real values on real-valued functions, and that $\ell = \ell_R + i\ell_I$ exhibits ℓ as a complex linear combination of continuous linear functionals taking real values on real-valued functions. This reduces the proposition to the case that ℓ takes real values on real-valued functions.

In this case, for $f \geq 0$, inspection gives the following: $\ell(f) = \ell^+(f) - \ell^-(f)$, $\ell^+(0) = \ell^-(0) = 0$, $\ell^+(cf) = c\ell^+(f)$ for $c \geq 0$, and $\ell^-(cf) = c\ell^-(f)$ for $c \geq 0$. In addition, $\ell^+(f) \geq 0$ for $f \geq 0$ because

$$\ell^+(f) = \sup_{0 \leq g \leq f} \ell(g) \geq f(0) = 0,$$

and $\ell^-(f) \geq 0$ for $f \geq 0$ because

$$\ell^-(f) = \ell^+(f) - \ell(f) = \sup_{0 \leq g \leq f} \ell(g) - \ell(f) \geq \ell(f) - \ell(f) = 0.$$

To complete the proof, all that we have to do is show that $\ell^+(f_1 + f_2) = \ell^+(f_1) + \ell^+(f_2)$ whenever $f_1 \geq 0$ and $f_2 \geq 0$. The argument for $\geq$ is that

$$\begin{aligned} \ell^+(f_1 + f_2) &= \sup_{0 \leq g \leq f_1 + f_2} \ell(g) \geq \sup_{\substack{g_1, g_2, \\ 0 \leq g_1 \leq f_1, \\ 0 \leq g_2 \leq f_2}} \ell(g_1 + g_2) \\ &= \sup_{0 \leq g_1 \leq f_1} \ell(g_1) + \sup_{0 \leq g_2 \leq f_2} \ell(g_2) = \ell^+(f_1) + \ell^+(f_2). \end{aligned}$$

For the reverse direction, let g be arbitrary with $0 \leq g \leq f_1 + f_2$, and set $g_1 = \min\{g, f_1\}$ and $g_2 = g - g_1$. Certainly $0 \leq g_1 \leq f_1$. Let us show that $0 \leq g_2 \leq f_2$. In fact,

$$\begin{aligned} g_2 = g - g_1 &= (g + f_1) - (f_1 + g_1) = \max\{g, f_1\} + \min\{g, f_1\} - (f_1 + g_1) \\ &= \max\{g, f_1\} + g_1 - (f_1 + g_1) = \max\{g, f_1\} - f_1. \end{aligned}$$

Thus g_2 is certainly ≥ 0. In addition, the computation

$$g_2 = \max\{g, f_1\} - f_1 \leq \max\{f_1 + f_2,\ f_1\} - f_1 = (f_1 + f_2) - f_1 = f_2$$

shows that g_2 is $\leq f_2$. Thus any g with $0 \leq g \leq f_1 + f_2$ gives us a corresponding decomposition

$$\begin{aligned} \ell(g) &= \ell(g_1 + g_2) = \ell(g_1) + \ell(g_2) \\ &\leq \sup_{0 \leq g_1 \leq f_1} \ell(g_1) + \sup_{0 \leq g_2 \leq f_2} \ell(g_2) = \ell^+(f_1) + \ell^+(f_2). \end{aligned}$$

Taking the supremum over g, we obtain $\ell^+(f_1 + f_2) \leq \ell^+(f_1) + \ell^+(f_2)$, and the proof is complete.

Let us reinterpret matters in terms of Borel measures. We begin with the real-valued case. Recall from Section IX.3 that a real-valued completely additive set function ρ on a σ-algebra is called a signed measure. It is bounded if $|\rho(E)| \le C$ for all E in the algebra. In this case Theorem 9.14 shows that it has a Jordan decomposition $\rho = \rho^+ - \rho^-$, where ρ^+ and ρ^- are uniquely determined finite measures such that any decomposition $\rho = \nu^+ - \nu^-$ as the difference of finite measures has $\rho^+ \le \nu^+$ and $\rho^- \le \nu^-$. We say that a bounded signed measure ρ on the Borel sets of the compact Hausdorff space X is a **regular Borel signed measure** if its Jordan decomposition is into regular Borel measures. If $\rho = \nu^+ - \nu^-$ is any decomposition of a bounded signed measure ρ on the Borel sets as the difference of regular Borel measures, then the equalities $\rho^+ \le \nu^+$ and $\rho^- \le \nu^-$ that compare the decomposition with the Jordan decomposition force ρ^+ and ρ^- to be regular, in view of Proposition 11.20. Hence ρ is a regular Borel signed measure.

The regular Borel signed measures form a real vector space $M(X, \mathbb{R})$. To see closure under vector space operations, we observe from the definition of regularity that the sum of two (nonnegative) regular Borel measures is a regular Borel measure. From this fact we can see that the sum of two regular Borel signed measures is regular and hence that $M(X, \mathbb{R})$ is closed under addition: in fact, if $\rho = \rho^+ - \rho^-$ and $\sigma = \sigma^+ - \sigma^-$ are given in their Jordan decompositions, then the formula $(\rho+\sigma)^+ - (\rho+\sigma)^- = (\rho^+ + \sigma^+) - (\rho^- + \sigma^-)$ shows that $\rho + \sigma$ is the difference of two regular Borel measures and hence is regular. Thus $M(X, \mathbb{R})$ is a real vector space.

Proposition 11.25. The real vector space $M(X, \mathbb{R})$ becomes a real normed linear space under the definition $\|\rho\| = \rho^+(X) + \rho^-(X)$, where $\rho = \rho^+ - \rho^-$ is the Jordan decomposition of ρ.

PROOF. Certainly $\|\rho\| \ge 0$ with equality if and only if $\rho = 0$. Also, if ρ has the Jordan decomposition $\rho = \rho^+ - \rho^-$, then $-\rho = \rho^- - \rho^+$ is the Jordan decomposition of $-\rho$, and it follows that $\|c\rho\| = |c|\|\rho\|$ for any real scalar c.

Finally consider $\|\rho + \sigma\|$. If $\rho = \rho^+ - \rho^-$ and $\sigma = \sigma^+ - \sigma^-$ are Jordan decompositions, then the formula $(\rho+\sigma)^+ - (\rho+\sigma)^- = (\rho^+ + \sigma^+) - (\rho^- + \sigma^-)$ shows that $(\rho + \sigma)^+ \le \rho^+ + \sigma^+$ and hence $(\rho + \sigma)^+(X) \le \rho^+(X) + \sigma^+(X)$. Similarly $(\rho + \sigma)^-(X) \le \rho^- X) + \sigma^-(X)$. Adding these inequalities, we obtain $\|\rho + \sigma\| \le \|\rho\| + \|\sigma\|$.

Returning to the statement of Proposition 11.24, let us write $C(X, \mathbb{R})$ or $C(X, \mathbb{C})$ for the space of continuous scalar-valued functions when the field of scalars is important, reserving the expression $C(X)$ for situations in which the scalars do not matter. Suppose that ℓ is a continuous linear function on $C(X)$ that takes real values on real-valued functions. The proposition shows that ℓ is the difference of two positive linear functionals. By Theorem 11.1, ℓ operates as the difference of two integrations: $\ell(f) = \int_X f\,d\nu^+ - \int_X f\,d\nu^-$, where ν^+ and ν^- are the regular Borel measures corresponding to ℓ^+ and ℓ^-. Then ℓ corresponds to a regular Borel signed measure ρ and is given by integration: $\ell(f) = \int_X f\,d\rho$, the integral with respect to the signed measure being interpreted as the difference of two integrals with respect to measures. Conversely any regular Borel signed measure ρ yields a continuous linear functional ℓ on $C(X)$ by the definition $\ell(f) = \int_X f\,d\rho$.

In particular the passage to integration gives us a real-linear mapping of $M(X, \mathbb{R})$ onto the space $C(X, \mathbb{R})^*$ of continuous linear functionals on the real vector space $C(X, \mathbb{R})$. Both of these spaces are normed linear spaces, and the theorem is that the map is one-one and that the norms match.

Theorem 11.26. The real-linear map of $M(X, \mathbb{R})$ onto $C(X, \mathbb{R})^*$ given by $\rho \mapsto \ell$ with $\ell(f) = \int_X f\,d\rho$ is one-one and norm preserving.

REMARK. As in Section V.9 the norm $\|\ell\|$ of ℓ is the least constant C such that $|\ell(f)| \le C\|f\|_{\sup}$ for all f. The constant C equals the supremum of $|\ell(f)|$ over all f with $\|f\|_{\sup} \le 1$.

PROOF. To see that the map is one-one, suppose that $\int_X f\,d\rho = 0$ for all f in $C(X, \mathbb{R})$. Then $\int_X f\,d\rho^+ = \int_X f\,d\rho^-$, and the uniqueness part of Theorem 11.1 shows that $\rho^+ = \rho^-$. Hence $\rho = \rho^+ - \rho^- = 0$.

Now suppose that ℓ and ρ correspond. Then we have

$$\begin{aligned} |\ell(f)| &= \left| \int_X f\,d\rho^+ - \int_X f\,d\rho^- \right| \\ &\le \int_X |f|\,d\rho^+ + \int_X |f|\,d\rho^- \\ &\le \rho^+(X)\|f\|_{\sup} + \rho^-(X)\|f\|_{\sup}. \end{aligned}$$

Taking the supremum over all f with $\|f\|_{\sup} \le 1$, we obtain

$$\|\ell\| \le \rho^+(X) + \rho^-(X) = \|\rho\|.$$

For the inequality in the reverse direction, let $\epsilon > 0$ be given, and let $X = P \cup N$ be a Hahn decomposition (Theorem 9.15) for ρ. By regularity of ρ^+ on P and ρ^- on N, choose compact subsets K_P and K_N with $K_P \subseteq P$, $K_N \subseteq N$, $\rho^+(P - K_P) < \epsilon$, and $\rho^-(N - K_N) < \epsilon$. Since $\rho^+(N) = 0$ and $\rho^-(P) = 0$,

$$\rho^+(X - K_P) < \epsilon \qquad \text{and} \qquad \rho^-(X - K_N) < \epsilon. \tag{$*$}$$

By Urysohn's Lemma (Corollary 10.43), we can find a continuous function $f : X \to [-1, 1]$ such that f is 1 on K_P and is -1 on K_N. Then

$$\begin{aligned} \big|\ell(f) - \|\rho\|\big| &\le \left| \int_{K_P} f\,d\rho - \|\rho^+\| \right| + \left| \int_{K_N} f\,d\rho - \|\rho^-\| \right| + \left| \int_{K_P^c \cap K_N^c} f\,d\rho \right| \\ &\le \big|\rho^+(K_P) - \rho^+(X)\big| + \big|\rho^-(K_N) - \rho^-(X)\big| + \left| \int_{K_P^c \cap K_N^c} f\,d\rho \right|. \end{aligned}$$

By $(*)$ the first two terms on the right side are each $< \epsilon$. Since $\rho^+(K_P^c \cap K_N^c) = \rho^+(P - K_P) < \epsilon$ and $\rho^-(K_P^c \cap K_N^c) = \rho^-(N - K_N) < \epsilon$, and since $\|f\|_{\sup} \le 1$, the third term on the right side is $\le 2\epsilon$. Therefore $\big|\ell(f) - \|\rho\|\big| < 4\epsilon$, and our function f has the property that $|\ell(f)| \ge (\|\rho\| - 4\epsilon)\|f\|_{\sup}$. In other words, $\|\ell\| \ge \|\rho\| - 4\epsilon$. Since ϵ is arbitrary, $\|\ell\| \ge \|\rho\|$. This completes the proof.

Now let us consider the case in which the values are complex. A **regular Borel complex measure** on the compact Hausdorff space X is an expression $\rho = \rho_R + i\rho_I$ in which ρ_R and ρ_I are regular Borel signed measures. In other words, it is a complex-valued set function whose real and imaginary parts are regular Borel signed measures. The space $M(X, \mathbb{C})$ of these is a complex vector space, and we shall make it into a normed linear space shortly. Meanwhile, the space $C(X, \mathbb{C})^*$ of continuous linear functionals on $C(X, \mathbb{C})$ is a complex normed linear space. Extending the definition of $\int_X f\,d\rho$ to handle members of $M(X, \mathbb{C})$, we see from Proposition 11.24 that the complex-linear map of $M(X, \mathbb{C})$ into $C(X, \mathbb{C})^*$ given by $\rho \mapsto \ell$ with $\ell(f) = \int_X f\,d\rho$ is one-one and onto.

To have a theorem in this case that parallels Theorem 11.26, we need to define the norm on $M(X, \mathbb{C})$. Doing so on an element ρ is not just a matter of combining the norms of the real and imaginary parts of ρ any more than writing the norm of a complex-valued L^1 function can be done in terms of the L^1 norms of the real and imaginary parts. A more subtle definition is needed.

We define the **total variation** $|\rho|$ of a member ρ of $M(X, \mathbb{C})$ to be the nonnegative set function whose value on a Borel set E is the supremum of all finite sums $\sum_{j=1}^n |\rho(E_j)|$ with $E = \bigcup_{j=1}^n E_j$ disjointly. The **total-variation norm** of the member ρ of $M(X, \mathbb{C})$ is defined to be $\|\rho\| = |\rho|(X)$. It is a simple matter to verify that the total-variation norm is indeed a norm.

Proposition 11.27. The total variation $|\rho|$ of a member ρ of $M(X, \mathbb{C})$ is a regular Borel measure, there exists a Borel function h with $\|h\|_{\sup} \le 1$ such that $\rho = h\,d|\rho|$, and the total-variation norm on $M(X, \mathbb{C})$ makes $M(X, \mathbb{C})$ into a normed linear space in such a way that $\big|\int_X f\,d\rho\big| \le \|\rho\|\|f\|_{\sup}$ for every bounded Borel function f. Moreover, $|\rho|$ equals $\rho^+ + \rho^-$ if ρ is real valued and has $\rho = \rho^+ - \rho^-$ as its Jordan decomposition.

REMARK. It follows that if ρ is real valued and if $X = P \cup N$ is a Hahn decomposition (Theorem 9.15) for ρ, then the corresponding function h may be taken to be $+1$ on P and $-N$ on N.

PROOF. To see that $|\rho|$ is additive, let E and F be disjoint Borel sets. If $E = \bigcup_{i=1}^m E_i$ disjointly and $F = \bigcup_{j=1}^n F_j$ disjointly, then $E \cup F = \big(\bigcup_{i=1}^m E_i\big) \cup \big(\bigcup_{j=1}^n F_j\big)$ disjointly, and hence $\sum_{i=1}^m |\rho(E_i)| + \sum_{j=1}^n |\rho(F_j)| \le |\rho|(E \cup F)$. Taking the supremum over systems $\{E_i\}$ and then over systems

$\{F_j\}$, we obtain $|\rho|(E) + |\rho|(F) \le |\rho|(E \cup F)$. In the reverse direction let $E \cup F = \bigcup_{k=1}^p G_k$ disjointly. Then $E = \bigcup_{k=1}^p (E \cap G_k)$ disjointly, and $F = \bigcup_{k=1}^p (F \cap G_k)$ disjointly. Hence

$$\begin{aligned}\sum_{k=1}^p &|\rho(G_k)| \\ &= \sum_{k=1}^p |\rho(E \cap G_k) + \rho(F \cap G_k)| \le \sum_{k=1}^p |\rho(E \cap G_k)| + \sum_{k=1}^p |\rho(F \cap G_k)|,\end{aligned}$$

and this is $\le |\rho|(E) + |\rho|(F)$. Taking the supremum over systems $\{G_k\}$, we obtain $|\rho|(E \cup F) \le |\rho|(E) + |\rho|(F)$. Thus $|\rho|$ is additive.

To prove that $|\rho|$ is completely additive, let $E = \bigcup_{n=1}^\infty E_n$ disjointly. For every N, $\sum_{n=1}^N |\rho|(E_n) = |\rho|(E_1 \cup \cdots \cup E_N) \le |\rho|(E)$, and hence $\sum_{n=1}^\infty |\rho|(E_n) \le |\rho|(E)$. For the reverse inequality let $\{G_k\}_{k=1}^p$ be a finite collection of disjoint Borel sets with union E. Then $E_n = \bigcup_{k=1}^p (E_n \cap G_k)$ disjointly, and hence

$$\begin{aligned}\sum_{k=1}^p |\rho(G_k)| = \sum_{k=1}^p |\rho(E \cap G_k)| &= \sum_{k=1}^p \Big| \sum_{n=1}^\infty \rho(E_n \cap G_k) \Big| \\ &\le \sum_{k=1}^p \sum_{n=1}^\infty |\rho(E_n \cap G_k)| = \sum_{n=1}^\infty \sum_{k=1}^p |\rho(E_n \cap G_k)| \le \sum_{n=1}^\infty |\rho|(E_n).\end{aligned}$$

Thus $|\rho|(E) \le \sum_{n=1}^\infty |\rho|(E_n)$, and $|\rho|$ is completely additive.

The measure $|\rho|$ is certainly finite on X and hence on all compact sets. To see regularity, we write $\rho = \rho_R + i\rho_I = \rho_R^+ - \rho_R^- + i\rho_I^+ - i\rho_I^-$. Writing a set E as the disjoint union of n sets E_i and writing out $\rho(E_i)$ according to this expansion of ρ, we see that $|\rho|(E) \le (\rho_R^+ + \rho_R^- + \rho_I^+ + \rho_I^-)(E)$. Each measure on the right side is regular, and Proposition 11.20 therefore shows that $|\rho|$ is regular.

For the existence of h, let us write ρ in terms of its real and imaginary parts as $\rho = \rho_R + i\rho_I$. If E is a Borel set, then the definitions give $|\rho|(E) \ge |\rho(E)| \ge |\rho_R(E)|$ and similarly $|\rho|(E) \ge |\rho_I(E)|$. Hence $\rho_R \ll |\rho|$ and $\rho_I \ll |\rho|$. By the Radon–Nikodym Theorem (Corollary 9.17), there exist functions h_R and h_I integrable $[d|\rho|]$ such that $\rho_R = h_R\, d|\rho|$ and $\rho_I = h_I\, d|\rho|$. Thus the $|\rho|$ integrable complex-valued function $h = h_R + ih_I$ has $\rho = h\, d|\rho|$. We shall show that h has $|h(x)| \le 1$ a.e. $[d|\rho|]$. If the contrary were the case, then there would exist a constant c with $|c| = 1$ and an $\epsilon > 0$ such that $\mathrm{Re}(ch) \ge 1+\epsilon$ on a set E of positive $|\rho|$ measure and we would have

$$\begin{aligned}\big|\textstyle\int_E h\, d|\rho|\big| = \big|\int_E ch\, d|\rho|\big| &\ge \mathrm{Re}\textstyle\int_E ch\, d|\rho| = \int_E \mathrm{Re}(ch)\, d|\rho| \\ &\ge (1+\epsilon)|\rho|(E) \ge (1+\epsilon)|\rho(E)| = (1+\epsilon)\big|\textstyle\int_E h\, d|\rho|\big|,\end{aligned}$$

a contradiction. Thus h exists as asserted.

The inequality $\left|\int_X f\,d\rho\right| \le \|\rho\|\|f\|_{\sup}$ follows from the existence of h since $\left|\int_X f\,d\rho\right| = \left|\int_X fh\,d|\rho|\right| \le \|fh\|_{\sup}\int_X d|\rho| \le \|f\|_{\sup}|\rho|(X) = \|f\|_{\sup}\|\rho\|$.

Finally if ρ is real valued, then any Borel set E satisfies $|\rho(E)| = |\rho^+(E) - \rho^-(E)| \le \rho^+(E) + \rho^-(E)$. If E is the disjoint union of Borel sets $E_1, \dots, E_n$, we consequently have

$$\sum_{j=1}^n |\rho(E \cap E_j)| \le \sum_{j=1}^n (\rho^+(E \cap E_j) + \rho^-(E \cap E_j)) = \rho^+(E) + \rho^-(E).$$

Taking the supremum over all decompositions of E of this kind gives $|\rho|(E) \le \rho^+(E) + \rho^-(E)$. For the reverse inequality let $X = P \cup N$ be a Hahn decomposition (Theorem 9.15) for ρ, so that $\rho^+(E) = \rho(P \cap E)$ and $\rho^-(E) = -\rho(N \cap E)$. Then E is the disjoint union of $E \cap P$ and $E \cap N$, and thus $\rho^+(E) + \rho^-(E) = |\rho(E \cap P)| + |\rho(E \cap N)| \le |\rho|(E)$. In other words, $|\rho| = \rho^+ + \rho^-$ as asserted.

Theorem 11.28. The one-one complex-linear map of $M(X, \mathbb{C})$ onto $C(X, \mathbb{C})^*$ given by $\rho \mapsto \ell$ with $\ell(f) = \int_X f\,d\rho$ is norm preserving.

PROOF. If f is in $C(X)$, then Proposition 11.27 gives $|\ell(f)| = \left|\int_X f\,d\rho\right| \le \|\rho\|\|f\|_{\sup}$. Taking the supremum over all f with $\|f\|_{\sup} \le 1$, we obtain $\|\ell\| \le \|\rho\|$.

For the reverse inequality let $\epsilon > 0$ be given, and choose a finite disjoint collection of Borel sets $E_1, \dots, E_n$ with union X such that $\sum_{i=1}^n |\rho(E_i)| \ge \|\rho\| - \epsilon$. Since $|\rho|$ is regular, we can find compact sets $K_i \subseteq E_i$ such that $|\rho|(E_i - K_i)| \le \epsilon/n$ for each i.

We shall define disjoint open sets U_i with $K_i \subseteq U_i$ for all i. First we find disjoint open sets U_1 and V_1 containing K_1 and $K_2 \cup \cdots \cup K_n$. Having inductively chosen disjoint open sets $U_1, \dots, U_j$ and V_j such that $K_i \subseteq U_i$ for $i \le j$ and $K_{i+1} \cup \cdots \cup K_n \subseteq V_i$, we use Corollary 10.22 to choose disjoint open subsets U_{i+1} and V_{i+1} of V_i containing K_{i+1} and $K_{i+2} \cup \cdots \cup K_n$. In this way we obtain the disjoint open sets $U_1, \dots, U_n$ with $K_i \subseteq U_i$ for all i.

For $1 \le i \le n$, choose $f_i \in C(X)$ with values in $[0, 1]$ such that f_i is 1 on K_i and is 0 off U_i. Choose $c_i \in \mathbb{C}$ for each i such that $c_i\rho(E_i) = |\rho(E_i)|$, and define $f_0 = \sum_{i=1}^n c_i f_i$. The function f_0 has $\|f_0\|_{\sup} = 1$ since the sets U_i are disjoint. Then

$$\begin{aligned}\ell(f_0) = \int_X f_0\,d\rho = \sum_{i=1}^n \int_{E_i} f_0\,d\rho &= \sum_{i=1}^n \Big(\int_{E_i} c_i\,d\rho + \int_{E_i} (f_0 - c_i)\,d\rho\Big)\\ &= \sum_{i=1}^n |\rho(E_i)| + \sum_{i=1}^n \int_{E_i - K_i} (f_0 - c_i)\,d\rho.\end{aligned}$$

Hence

$$\Big|\ell(f_0) - \sum_{i=1}^{n} |\rho(E_i)|\Big| \le \sum_{i=1}^{n} \int_{E_i - K_i} |f_0 - c_i|\, d|\rho|$$
$$\le 2 \sum_{i=1}^{n} |\rho|(E_i - K_i)| \le 2 \sum_{i=1}^{n} \epsilon/n = 2\epsilon$$

and

$$\Big|\ell(f_0) - \|\rho\|\Big| \le \Big|\ell(f_0) - \sum_{i=1}^{n} |\rho(E_i)|\Big| + \Big|\sum_{i=1}^{n} |\rho(E_i)| - \|\rho\|\Big| \le 3\epsilon.$$

Therefore

$$\|\ell\| = \|\ell\|\|f_0\|_{\sup} \ge |\ell(f_0)| \ge \|\rho\| - \Big|\ell(f_0) - \|\rho\|\Big| \ge \|\rho\| - 3\epsilon.$$

Since ϵ is arbitrary, $\|\ell\| \ge \|\rho\|$.

5. Problems

In all problems for this chapter, X is assumed to be a locally compact Hausdorff space. Sometimes additional hypotheses are imposed on X.

1. (a) Prove that if X is σ-compact, then the σ-algebra of Borel subsets of X coincides with the σ-algebra of intersections of X with the Borel subsets of the one-point compactification X^*.
   (b) Prove that if X is an uncountable discrete space, then the σ-algebra of Borel subsets of X is strictly smaller than the σ-algebra of intersections of X with the Borel subsets of the one-point compactification X^*.
2. Prove that if X is σ-compact and $f : X \to \mathbb{C}$ is continuous, then f is a Borel function.
3. Suppose that X is σ-compact. Prove that if μ is a regular Borel measure on X and if f is Borel measurable, then there exists a Baire measurable function g such that $f = g$ except on a Borel set of μ measure 0.
4. **(Lusin's Theorem)** Let X be compact, let μ be a regular Borel measure on X, let f be a Borel function on X, and let $\epsilon > 0$ be given. By first considering simple functions and then passing to the limit via Egoroff's Theorem, prove that there exists a compact subset K of X with $\mu(K^c) < \epsilon$ such that $f\big|_K$ is continuous.
5. This problem establishes the rotation invariance of the Borel measure $d\omega$ on the sphere $S^2 \subseteq \mathbb{R}^3$ obtained from Riemann integration with respect to $\sin\theta_1\, d\theta_1\, d\theta_2$, where θ_1 and θ_2 are latitude and longitude with $0 \le \theta_1 \le \pi$ and $0 \le \theta_2 \le 2\pi$. The measure $d\omega$ was constructed by means of the Riesz Representation Theorem as one of the examples in Section 2.

(a) A rotation in $\mathbb{R}^3$ is the linear function L determined by a matrix A with $AA^{\mathrm{tr}} = 1$ and $\det A = 1$. For $0 < a < 1 < b < \infty$, let S_{ab} be the subset of $\mathbb{R}^3$ given in spherical coordinates by $a < r < b, 0 \le \theta_1 \le \pi, 0 \le \theta_2 \le 2\pi$. Show that S_{ab} is carried to itself by any such rotation L.

(b) For any bounded Borel function $F : S_{ab} \to \mathbb{C}$, let $(LF)(x) = F(L^{-1}x)$ if x is in S_{ab} and L is a rotation. Prove that $\int_{S_{ab}} LF\,dx = \int_{S_{ab}} F\,dx$.

(c) Let $f : S^2 \to \mathbb{C}$ be any continuous function, and define $(Lf)(\omega) = f(L^{-1}\omega)$. Extend f to a function F defined on S_{ab} by the definition $F(r\omega) = f(\omega)$. Prove that $\int_{S_{ab}} F\,dx = \big(\int_a^b r^2\,dr\big)\big(\int_{S^2} f(\omega)\,d\omega\big)$ and deduce that $\int_{S^2} Lf\,d\omega = \int_{S^2} f\,d\omega$.

(d) Deduce from (c) that $d\omega(L(E)) = d\omega(E)$ for every Borel subset E of S^2.

6. Let X be compact.

(a) Let $\{K_\alpha\}$ be a collection of compact subsets of X closed under finite intersections, and let $K = \bigcap_\alpha K_\alpha$. Prove that every regular Borel measure μ on X has the property that $\mu(K) = \inf_\alpha \mu(K_\alpha)$.

(b) If μ is a nonzero regular Borel measure on X assuming only the values 0 and 1, prove that μ is a point mass.

(c) If μ is a nonzero regular Borel measure on X with

$$\int_X fg\,d\mu = \Big(\int_X f\,d\mu\Big)\Big(\int_X g\,d\mu\Big)$$

for all f and g in $C(X)$, prove that μ is a point mass.

(d) If ℓ is a positive linear functional on $C(X)$ that is multiplicative in the sense that $\ell(fg) = \ell(f)\ell(g)$ for all f and g in $C(X)$, prove that ℓ is zero or ℓ is evaluation at some point of X.

7. This problem continues the investigation of harmonic functions and Poisson integrals in the unit disk of $\mathbb{R}^2$, following up on Problems 7–8 at the end of Chapter IX. Problem 8 in that series provides orientation. The new ingredient for the present problem is weak-star convergence of sequences in $M(S^1, \mathbb{C})$ against $C(S^1)$, where S^1 is the unit circle.

(a) State and prove a characterization of the harmonic functions $u(r, \theta)$ on the open unit disk such that $\sup_{0 \le r < 1} \|u(r,\,\cdot\,)\|_1$ is finite.

(b) **(Herglotz's Theorem)** Prove that if $u(r, \theta)$ is a nonnegative harmonic function on the open unit disk, then there is a Borel measure μ on the circle such that $u(r, \theta) = \int_{(-\pi,\pi]} P_r(\theta - \varphi)\,d\mu(\varphi)$.

Problems 8–10 construct a Borel measure μ on a compact space such that μ is not regular. The totally ordered set Ω of countable ordinals was introduced in Problems 25–33 at the end of Chapter V. Let $\Omega^* = \Omega \cup \{\infty\}$, totally ordered so that every element of Ω is less than $\{\infty\}$. Give Ω^* the order topology, as discussed in Problems 25–32 at the end of Chapter X.

8. Prove that Ω^* is compact Hausdorff.
9. Prove that the class of all relatively closed uncountable subsets of Ω is closed under the formation of countable intersections.
10. Define μ on the Borel sets of Ω^* to be 1 on those sets E such that $E - \{\infty\}$ contains a relatively closed uncountable subset of Ω, and put $\nu(E) = 0$ otherwise. Prove that μ is a Borel measure that is not regular.

Problems 11–14 concern decomposing any finite Borel measure on a compact X into a regular Borel measure and a "purely irregular" Borel measure. They make use of Zorn's Lemma (Section A9 of the appendix). A Borel measure μ will be called **purely irregular** if there is no nonzero regular Borel measure ν such that $0 \le \nu(E) \le \mu(E)$ for every Borel set E.

11. Use Zorn's Lemma to show that any Borel measure on X is the sum of a regular Borel measure and a purely irregular Borel measure.
12. Prove that if ν is a regular Borel measure, if μ is purely irregular, and if $0 \le \mu \le \nu$, then $\mu = 0$.
13. Deduce from the Jordan decomposition (Theorem 9.14) that the decomposition of Problem 11 is unique.
14. Prove that the irregular Borel measure constructed in Problem 10 is purely irregular.

Problems 15–19 concern extension of measures from finite products of compact metric spaces to countably infinite such products. Let X be a compact metric space, and for each integer $n \ge 1$, let X_n be a copy of X. Define $\Omega^{(N)} = \bigtimes_{n=1}^{N} X_n$, and let $\Omega = \bigtimes_{n=1}^{\infty} X_n$. Each of $\Omega^{(N)}$ and Ω is given the product topology. If E is a Borel subset of $\Omega^{(N)}$, we can regard E as a subset of Ω by identifying E with $E \times \big(\bigtimes_{n=N+1}^{\infty} X_n\big)$. In this way any Borel measure on $\Omega^{(N)}$ can be regarded as a measure on a certain σ-subalgebra $\mathcal{F}_n$ of $\mathcal{B}(\Omega)$.

15. Prove that $\bigcup_{n=1}^{\infty} \mathcal{F}_n = \mathcal{F}$ is an algebra.
16. Let ν_n be a (regular) Borel measure on $\Omega^{(n)}$ with $\nu(\Omega^{(n)}) = 1$, and regard ν_n as defined on $\mathcal{F}_n$. Suppose for each n that ν_n agrees with ν_{n+1} on $\mathcal{F}_n$. Define $\nu(E)$ for E in $\mathcal{F}$ to be the common value of $\nu_n(E)$ for n large. Prove that ν is nonnegative additive, and prove that in a suitable sense ν is regular on $\mathcal{F}$.
17. Using the kind of regularity established in the previous problem, prove that ν is completely additive on $\mathcal{F}$.
18. In view of Problems 16 and 17, ν extends to a measure on the smallest σ-algebra for Ω containing $\mathcal{F}$. Prove that this σ-algebra is $\mathcal{B}(\Omega)$.
19. Let X be a 2-point space, and let ν_n be 2^{-n} on each one-point subset of $\Omega^{(n)}$. Exhibit a homeomorphism of Ω onto the standard Cantor set in $[0, 1]$ that carries ν to the Cantor measure defined in Problems 17–20 at the end of Chapter VI.

CHAPTER XII

Hilbert and Banach Spaces

Abstract. This chapter develops the beginnings of abstract functional analysis, a subject designed to study properties of functions by treating the functions as the members of a space and formulating the properties as properties of the space.

Section 1 defines Banach spaces as complete normed linear spaces and gives a number of examples of these. The space of bounded linear operators from one normed linear space to another is a normed linear space, and it is a Banach space if the range is a Banach space.

Sections 2–3 concern Hilbert spaces. These are Banach spaces whose norms are induced by inner products. Section 2 shows that closed vector subspaces of such a space have orthogonal complements, and it shows the role of orthonormal bases for such a space. Section 3 concentrates on bounded linear operators from a Hilbert space to itself and constructs the adjoint of each such operator.

Sections 4–6 prove the three main abstract theorems about the norm topology of general normed linear spaces—the Hahn–Banach Theorem, the Uniform Boundedness or Banach–Steinhaus Theorem, and the Interior Mapping Principle. A number of consequences of these theorems are given. The second and third of the theorems require some hypothesis of completeness.

1. Definitions and Examples

Functional analysis puts into practice an idea from the early twentieth century, that sometimes properties of functions become clearer when the functions are regarded as the members of a space and the properties are formulated as properties of the space. We encountered some simple examples of this situation already in Chapter II in the examples of metric spaces. Uniform convergence was encoded in the metric on spaces of functions, and other kinds of convergence were captured by other metrics. In Chapter V we introduced the spaces $L^1(X)$, $L^2(X)$, and $L^\infty(X)$ of functions (or really equivalence classes of functions), all of which were proved to be complete. The property of completeness was a useful property of the space as a whole that led, for one thing, to the Riesz–Fischer Theorem in Chapter VI. More complicated properties led us to various kinds of differentiability of integrals in $\mathbb{R}^n$ in Chapters VI and IX and to boundedness of the Hilbert transform in Chapter IX. The development of measure theory on locally compact Hausdorff spaces in Chapter XI rested on an analysis of positive linear functionals on the space of continuous functions of compact support.

The different spaces—of functions, measures, and whatever else—that arise in this way have some properties in common, and we study them in this chapter in a setting that emphasizes these common properties. We shall work with normed linear spaces, which were defined in Section V.9. With such spaces the field of scalars $\mathbb{F}$ can be either $\mathbb{R}$ or $\mathbb{C}$. Recall then that a **normed linear space** X is a vector space over $\mathbb{F}$ with a **norm**, i.e., a function $\|\cdot\|$ from X to $[0, +\infty)$ such that $\|x\| \geq 0$ with equality if and only if $x = 0$, $\|cx\| = |c|\|x\|$ if c is a scalar, and $\|x + y\| \leq \|x\| + \|y\|$. The norm yields a metric $d(x, y) = \|x - y\|$, and we can then speak of the norm topology on X. Proposition 5.55 showed that addition and scalar multiplication are continuous, that the closure of any vector subspace of X is a vector subspace, and that the set of all finite linear combinations of members of a subset S of X is dense in the smallest closed subspace containing S.

Completeness plays an increasingly important role as one studies such spaces, and it is customary to introduce a definition to incorporate this notion: a normed linear space X is a **Banach space** if X is complete as a metric space. The metric-space completion of a normed linear space is automatically a normed linear space that is complete, hence is a Banach space.

Let us consider some examples of normed linear spaces, some old and some new. Except as indicated, they will all be Banach spaces.

EXAMPLES.

(1) Euclidean space $\mathbb{R}^n$ and complex Euclidean space $\mathbb{C}^n$, written briefly as $\mathbb{F}^n$. The space consists of n-tuples of scalars $a = (a_1, \dots, a_n)$ with $\|a\|$ equal to the Euclidean norm $|a|$ of Section II.1, namely $\|a\| = \left(\sum_{k=1}^n |a_k|\right)^{1/2}$. It was remarked in Section II.7 that these spaces are complete, hence are Banach spaces.

(2) Finite-dimensional normed linear spaces. It can be shown that each finite-dimensional normed linear space X is complete. In fact, any linear map carrying a vector-space basis of X to a vector-space basis of some $\mathbb{F}^n$, normed as in the previous example, can be shown to be uniformly continuous with a uniformly continuous inverse, and the completeness of X follows.

(3) $B(S)$, the space of bounded scalar-valued functions on a nonempty set S with the supremum norm, defined in Section II.1. Proposition 2.44 establishes the completeness.

(4) $C(S)$, the space of bounded continuous scalar-valued functions on a metric space or topological space S, defined in Section II.4 in the metric case and Section X.5 in general. The norm is the supremum norm. Corollary 2.45 and Proposition 10.30 establish the completeness of $C(S)$. When S is locally compact Hausdorff, we defined $C_0(S)$ in Section XI.4 to be the subspace of $C(S)$ of all members vanishing at infinity. This is complete. However, the subspace $C_{\text{com}}(S)$ of continuous scalar-valued functions of compact support is usually not complete.

(5) $L^p(S, \mathcal{A}, \mu)$, the space of equivalence classes of p^{th}-power integrable functions on a measure space $(S, \mathcal{A}, \mu)$. This is a normed linear space for $1 \leq p < \infty$ with norm $\|f\|_p = \big(\int_S |f(s)|^p \, d\mu(s)\big)^{1/p}$. These spaces were introduced in Section V.9 for $p = 1$ and $p = 2$ and in Section IX.1 for general p. Theorem 5.59 established the completeness for $p = 1$ and $p = 2$, and Theorem 9.6 established the completeness for general p.

(6) $L^\infty(S, \mathcal{A}, \mu)$, the space of equivalence classes of essentially bounded functions on a measure space $(S, \mathcal{A}, \mu)$. This is a normed linear space with norm the essential supremum norm. This space was introduced in Section V.9 and was proved to be complete in Theorem 5.59.

(7) Sequence spaces c, c_0, and ℓ_n^p and ℓ^p for $1 \leq p \leq \infty$. These are special cases of various examples above. The space ℓ_n^p is $L^p(S, \mathcal{A}, \mu)$ when $S = \{1, 2, \ldots, n\}$, $\mathcal{A}$ is the set of all subsets, and μ is counting measure, the norm being $\|(a_1, \ldots, a_n)\| = \big(\sum_{k=1}^n |a_k|^p\big)^{1/p}$ if $p < \infty$ and being $\|(a_1, \ldots, a_n)\| = \max_{1 \leq k \leq n} |a_k|$ if $p = \infty$. The space ℓ_n^p specializes to $\mathbb{F}^n$ when $p = 2$. The space ℓ^p is the version of ℓ_n^p when S is the set of positive integers; the members of this space are thus all sequences for which the norm is finite. The sequence spaces c and c_0 can be regarded as subspaces of $C(S)$ when S is the set of positive integers. The space c consists of all convergent sequences, and c_0 is the space of sequences vanishing at infinity; in both cases the norm is the supremum norm. All these examples are Banach spaces. They tend to be useful in testing guesses about properties of normed linear spaces. We shall not need them explicitly, and this traditional notation for them will not recur after the end of this section.

(8) $M(S)$, S being a compact Hausdorff space. This is the space of regular Borel signed or complex measures on S, introduced as $M(S, \mathbb{R})$ or $M(S, \mathbb{C})$ in Section XI.4. The norm is the total-variation norm. Theorems 11.26 and 11.28 identify these spaces with duals of spaces of continuous functions, and Proposition 12.1 below will show that they are complete as a consequence.

(9) $C^N([a, b])$, the space of scalar-valued functions on a bounded interval $[a, b]$ with N bounded derivatives, the norm being

$$\|f\| = \sum_{j=1}^{N} \sup_{a \leq s \leq b} |f^{(j)}(s)|.$$

It is shown in Problem 2 at the end of the chapter that this space is complete. This space is an indication of how normed linear spaces can carry information about derivatives. Indeed, normed linear spaces carrying information about derivatives play a significant role in the subject of partial differential equations.[1]

[1]This is one of the themes of the companion volume *Advanced Real Analysis*.

(10) $H^\infty(D)$, the space of bounded functions in the open unit disk $D = \{|z| < 1\}$ in $\mathbb{C}$ such that the function is given by a convergent power series. The norm is the supremum norm. It is shown in Problem 3 at the end of the chapter that this space is complete.

(11) $A(D)$, the space of bounded continuous functions on the closed unit disk whose restriction to the open unit disk is given by a convergent power series. The norm is the supremum norm. It is shown in Problem 3 at the end of the chapter that this space is complete.

Two further kinds of normed linear spaces are worth mentioning now. One is that any real or complex inner-product space X in the sense of Section II.1 gives an example of a normed linear space. Recall that an **inner product** on X is a function $(\cdot\,,\,\cdot)$ from $X \times X$ to $\mathbb{F}$ that is linear in the first variable, is conjugate linear in the second variable, is symmetric if $\mathbb{F} = \mathbb{R}$ or Hermitian symmetric if $\mathbb{F} = \mathbb{C}$, and has $(x, x) \geq 0$ for all x with equality if and only if $x = 0$. Such an inner product satisfies the Schwarz inequality $|(x, y)| \leq (x, x)^{1/2}(y, y)^{1/2}$, according to Lemma 2.2, and then the definition $\|x\| = (x, x)^{1/2}$ makes X into a normed linear space, according to Proposition 2.3.

As a normed linear space, an inner-product space may or may not be complete. Any space $L^2(S, \mathcal{A}, \mu)$, with $(f, g) = \int_S f\bar{g}\,d\mu$, is an example in which the associated normed linear space is complete. An inner-product space whose associated normed linear space is complete is called a **Hilbert space**.

The other kind of normed linear space worth mentioning now involves bounded linear operators. Recall from Section V.9 that a linear function $L : X \to Y$ between two normed linear spaces with respective norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ is often called a **linear operator**. Proposition 5.57 showed that a linear operator L is continuous at a point if and only if it is continuous everywhere, if and only if it is uniformly continuous, if and only if it is **bounded** in the sense that $\|L(x)\|_Y \leq M\|x\|_X$ for some constant M and all x in X. The least such constant M is called the **operator norm** of L, written $\|L\|$. We can define addition and scalar multiplication on bounded linear operators from X to Y by addition and scalar multiplication of their values:

$$(L_1 + L_2)(x) = L_1(x) + L_2(x) \qquad \text{and} \qquad (cL)(x) = cL(x).$$

Then $L_1 + L_2$ and cL are linear operators by the elementary theory of vector spaces, and the inequalities

$$\begin{aligned}\|(L_1 + L_2)(x)\|_Y &= \|L_1(x) + L_2(x)\|_Y \leq \|L_1(x)\|_Y + \|L_2(x)\|_Y \\ &\leq \|L_1\|\|x\|_X + \|L_2\|\|x\|_X = (\|L_1\| + \|L_2\|)\|x\|_X\end{aligned}$$

and $$\|(cL)(x)\|_Y = \|cL(x)\|_Y = |c|\|L(x)\|_Y \leq |c|\|L\|\|x\|_X$$

show that $L_1 + L_2$ and cL are bounded with $\|L_1 + L_2\| \le \|L_1\| + \|L_2\|$ and $\|cL\| \le |c|\|L\|$. Applying the latter conclusion to c^{-1} when $c \ne 0$ gives $\|L\| = \|c^{-1}(cL)\| \le |c|^{-1}\|cL\| \le |c|^{-1}|c|\|L\| = \|L\|$, and we conclude that $\|cL\| = |c|\|L\|$. Since it is plain that $\|L\| \ge 0$ with equality if and only if $L = 0$, the set of bounded linear operators from X to Y, with the operator norm, is a normed linear space. We denote this normed linear space by $\mathcal{B}(X, Y)$.

Proposition 12.1. If X and Y are normed linear spaces and if Y is complete, then the normed linear space $\mathcal{B}(X, Y)$ is a Banach space.

REMARKS. In the special case in which Y is the set $\mathbb{F}$ of scalars, the linear operators are called **linear functionals**, in terminology we have used repeatedly. The normed linear space $\mathbb{F} = \mathbb{F}^1$ is complete, and therefore the normed linear space of bounded linear functionals on X is a Banach space. The space of bounded linear functionals is called the **dual space** of X and is denoted by X^*. More explicitly the norm of an element x^* of X^* is[2]

$$\|x^*\| = \sup_{\|x\| \le 1} |x^*(x)|.$$

Proposition 12.1 is implicitly saying that X^* is always complete.

PROOF. Let $\{L_n\}$ be a Cauchy sequence in $\mathcal{B}(X, Y)$. Since in any metric space the members of a Cauchy sequence are at a bounded distance from any particular element, the sequence $\{\|L_n\|\}$ is bounded. Let $C = \sup_n \|L_n\|$.

If x is in X, then $\{L_n(x)\}$ is a Cauchy sequence since $\|L_m(x) - L_n(x)\|_Y \le \|L_m - L_n\|\|x\|_X$. By completeness of Y, $L(x) = \lim_n L_n(x)$ exists. Continuity of addition and scalar multiplication in X implies that $L(x+x') = \lim_n L_n(x+x') = \lim_n(L_n(x) + L_n(x')) = \lim_n L_n(x) + \lim_n L_n(x') = L(x) + L(x')$ and that $L(cx) = \lim_n L_n(cx) = \lim_n(cL_n(x)) = c\lim_n L_n(x) = cL(x)$. Therefore L is a linear operator.

For boundedness of L, we have $\|L_n(x)\|_Y \le \|L_n\|\|x\|_X \le C\|x\|_X$ for all n. Hence continuity of the norm function implies that $\|L(x)\|_Y = \|\lim L_n(x)\|_Y \le \liminf_n \|L_n(x)\|_Y \le C\|x\|_X$, and L is bounded with $\|L\| \le C$.

To complete the proof, we show that $\|L_n - L\| \to 0$. Assuming the contrary, we can pass to a subsequence and then change notation so that $\|L_n - L\| \ge \epsilon$ for some $\epsilon > 0$ for all n. Then for each n, we can find x_n in X with $\|x_n\|_X = 1$

[2]A superscript * has also been used in this book to indicate a one-point compactification, but there need never be any confusion about this notation. One-point compactifications arise in practice only for locally compact Hausdorff spaces, and one can show that a normed linear space is locally compact only if it is finite dimensional, For finite-dimensional normed linear spaces it is always clear from the context whether * refers to the dual space or to the one-point compactification.

such that $\|L_n(x_n) - L(x_n)\|_Y \geq \epsilon/2$. Choose and fix N so that $m \geq N$ implies $\|L_N - L_m\| \leq \epsilon/4$. Whenever $m \geq N$, the triangle inequality gives

$$\begin{aligned}\|L_m(x_N) - L(x_N)\|_Y &\geq \|L_N(x_N) - L(x_N)\|_Y - \|L_N(x_N) - L_m(x_N)\|_Y \\ &\geq \tfrac{\epsilon}{2} - \|L_N - L_m\|\|x_N\|_X = \tfrac{\epsilon}{2} - \|L_N - L_m\| \geq \tfrac{\epsilon}{4},\end{aligned}$$

in contradiction to the fact that $\lim_m L_m(x_N) = L(x_N)$.

EXAMPLES OF DUAL SPACES.

(1) $L^p(S, \mathcal{A}, \mu)^* \cong L^{p'}(S, \mathcal{A}, \mu)$ if $1 \leq p < \infty$, μ is σ-finite, and p' is the **dual index** with $\frac{1}{p} + \frac{1}{p'} = 1$, according to the Riesz Representation Theorem (Theorem 9.19). Specifically to each x^* in $L^p(S, \mathcal{A}, \mu)^*$ corresponds a unique g in $L^{p'}(S, \mathcal{A}, \mu)$ with $x^*(f) = \int_S fg\, d\mu$ for all f in $L^p(S, \mathcal{A}, \mu)$, and this g has $\|x^*\| = \|g\|_{p'}$. It can be shown that the hypothesis of σ-finiteness of μ can be dropped if $1 < p < \infty$, but Problem 4 at the end of Chapter IX shows that the hypothesis cannot be completely dropped for $p = 1$.

(2) $(\ell_n^p)^* \cong \ell_n^{p'}$ and $(\ell^p)^* \cong \ell^{p'}$ for $1 \leq p < \infty$ if p' is the dual index. This is a special case of Example 1. In particular, the first of these duality results for $p = 2$ says that $(\mathbb{R}^n)^* \cong \mathbb{R}^n$ and $(\mathbb{C}^n)^* \cong \mathbb{C}^n$.

(3) $C(S)^* \cong M(S)$ if S is a compact Hausdorff space, according to Theorems 11.26 and 11.28. Specifically to each x^* in $C(S)^*$ corresponds a unique ρ in $M(S)$ with $x^*(f) = \int_S f\, d\rho$ for all f in $C(S)$, and this ρ has $\|x^*\| = \|\rho\|$. Since $M(S)$ is in this way identified as the dual space of some normed linear space, it follows from Proposition 12.1 that $M(S)$ is a Banach space.

(4) $(\ell_n^\infty)^* \cong \ell_n^1$ and $(c_0)^* \cong \ell^1$. The isomorphism $(\ell_n^\infty)^* \cong \ell_n^1$ is the special case of Example 3 in which $S = \{1, \ldots, n\}$. To see the isomorphism $(c_0)^* \cong \ell^1$, we take S to be the set of positive integers and form the one-point compactification S^*. The continuous scalar-valued functions on S^*, with their supremum norm, can be identified with the normed linear space c of convergent sequences. Thus Example 3 in this setting says that $c^* \cong M(S^*)$. The members of c_0 are the members of c that vanish at ∞, and any point mass at ∞ in a member of $M(S^*)$ has no effect on the subspace c_0. It readily follows that the dual of c_0 consists of the members of $M(S^*)$ with no point mass at ∞, and these elements, with their norm, may be identified with ℓ^1.

From one point of view, Hilbert spaces are particularly simple Banach spaces, and we shall study them first. The geometry of Hilbert space will be the topic of the next section, and the section after that will give a brief introduction to bounded linear operators from a Hilbert space to itself.

2. Geometry of Hilbert Space

Hilbert spaces were defined in Section 1 as complete normed linear spaces whose norms arise from an inner product. Euclidean space $\mathbb{R}^n$ and complex Euclidean space $\mathbb{C}^n$ are examples, and every space $L^2(S, \mathcal{A}, \mu)$ with $(f, g) = \int_S f\bar{g}\,d\mu$ is a Hilbert space. We shall see in this section that every Hilbert space shares many geometric facts in common with the finite-dimensional examples $\mathbb{R}^n$ and $\mathbb{C}^n$. The expansion of square integrable functions on $[-\pi, \pi]$ in Fourier series will be seen to be an example of expansion of all members of a Hilbert space in terms of an "orthonormal basis."

Let H be a real or complex Hilbert space with inner product $(\,\cdot\,,\,\cdot\,)$ and with norm $\|\cdot\|$ given by $\|u\| = (u, u)^{1/2}$. Lemma 2.2 shows that H satisfies the **Schwarz inequality**

$$|(u, v)| \le \|u\|\|v\| \qquad \text{for all } u \text{ and } v \text{ in } H.$$

The Schwarz inequality implies the estimate

$$|(u, v)-(u_0, v_0)| \le |(u-u_0, v)|+|(u_0, v-v_0)| \le \|u-u_0\|\|v\|+\|u_0\|\|v-v_0\|,$$

from which it follows that the inner product is a continuous function of two variables.

We shall make frequent use of the formula

$$\|u + v\|^2 = \|u\|^2 + 2\operatorname{Re}(u, v) + \|v\|^2,$$

which is what one combines with the Schwarz inequality to prove the triangle inequality for the norm. With the additional hypothesis that $(u, v) = 0$, this formula reduces to the **Pythagorean Theorem**

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Direct expansion of the norms squared in terms of the inner product shows that H satisfies the **parallelogram law**

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2 \qquad \text{for all } u \text{ and } v \text{ in } H.$$

Actually, there is a converse to this formula, due to Jordan and von Neumann, whose details are left to Problems 19–24 at the end of the chapter: a Banach space is a Hilbert space if its norm satisfies the parallelogram law. The idea is that the inner product in a Hilbert space can be computed from the identity

$$(u, v) = \frac{1}{4} \sum_k i^k \, \|u + i^k v\|^2,$$

where the sum extends for $k \in \{0, 2\}$ if the scalars are real and extends for $k \in \{0, 1, 2, 3\}$ if the scalars are complex. This identity goes under the name **polarization**. For the result of Jordan and von Neumann, one *defines* (u, v) by this formula, shows that the result is an inner product, and proves that $\|u\|^2 = (u, u)$.

The following lemma, which makes use of the completeness, is the key to all the geometry.

Lemma 12.2. If M is a closed vector subspace of the Hilbert space H and if u is in H, then there is a vector v in M with

$$\|u - v\| = \inf_{w \in M} \|u - w\|.$$

REMARK. Examination of the proof will show that we do not make full use of the assumption that M is closed under addition and scalar multiplication, only that M is closed under passage to convex combinations, i.e., that x and y in M imply that $tx + (1 - t)y$ is in M for all t with $0 \le t \le 1$. Thus it is enough to assume that M is a closed convex set, not necessarily a closed vector subspace.

PROOF. Let $d = \inf_{w \in M} \|u - w\|$, and choose a sequence $\{w_n\}$ in M with $\|u - w_n\| \to d$. By the parallelogram law,

$$\|2u - (w_n + w_m)\|^2 + \|w_n - w_m\|^2 = 2(\|u - w_m\|^2 + \|u - w_n\|^2) \longrightarrow 4d^2.$$

Since $\frac{1}{2}(w_n + w_m)$ is in M,

$$\|2u - (w_n + w_m)\|^2 = 4\|u - \tfrac{1}{2}(w_n + w_m)\|^2 \ge 4d^2.$$

We conclude that $\|w_n - w_m\|^2 \to 0$, and $\{w_n\}$ is Cauchy. By completeness of H, $\{w_n\}$ is convergent. If $v = \lim w_n$, then v is in M since M is topologically closed. Since $\|u - w_n\| \to d$, continuity of the norm gives $\|u - v\| = d$.

Two vectors u and v in H are said to be **orthogonal** if $(u, v) = 0$. The set of all vectors orthogonal to a subset M of H is denoted by $M^\perp$. In symbols,

$$M^\perp = \{u \in H \mid (u, v) = 0 \text{ for all } v \in M\}.$$

We see by inspection that $M^\perp$ is a closed vector subspace. Moreover, $M \cap M^\perp = 0$ since any u in $M \cap M^\perp$ must have $(u, u) = 0$. The subspace $M^\perp$ will be of greatest interest when M is a closed vector subspace, as a consequence of the following proposition.

Proposition 12.3 (Projection Theorem). If M is a closed vector subspace of the Hilbert space H, then every u in H decomposes uniquely as $u = v + w$ with v in M and w in $M^\perp$.

REMARKS. One writes $H = M \oplus M^\perp$ to express this unique decomposition of vector spaces. Because of this proposition, $M^\perp$ is often called the **orthogonal complement** of the closed vector subspace M. It is essential that M be closed in this proposition. In fact, consider the vector subspace M of polynomials in $L^2([0,1])$. This is dense as a consequence of the Weierstrass Approximation Theorem, and consequently no L^2 function other than 0 can be in $M^\perp$. Thus not every member of L^2 is the sum of a member of M and a member of $M^\perp$.

PROOF. Uniqueness follows from the fact that $M \cap M^\perp = 0$. For existence let u be in H, and choose v in M by Lemma 12.2 with $\|u - v\| = \inf_{w \in M} \|u - w\|$. If m is any member of M with $\|m\| = 1$, then the vector $v + (u - v, m)m$ is in M and the formula $\|x - y\|^2 = \|x\|^2 - 2\,\mathrm{Re}(x, y) + \|y\|^2$ gives

$$\begin{aligned} \|u - v\|^2 &\le \|u - v - (u - v, m)m\|^2 \\ &= \|u - v\|^2 - 2|(u - v, m)|^2 + |(u - v, m)|^2 \\ &= \|u - v\|^2 - |(u - v, m)|^2. \end{aligned}$$

Hence $(u - v, m) = 0$. Since every nonzero member of M is a scalar multiple of a member with $\|m\| = 1$, $u - v$ is in $M^\perp$.

Corollary 12.4. If M is a closed vector subspace of the Hilbert space H, then $M^{\perp\perp} = M$.

PROOF. From the definition we see that $M \subseteq M^{\perp\perp}$. If u is in $M^{\perp\perp}$, write $u = m + m^\perp$ with $m \in M$ and $m^\perp \in M^\perp$ by Proposition 12.3. Then $0 = m^\perp + (m - u)$ with $m^\perp \in M^\perp$ and $m - u \in M^{\perp\perp}$. By the uniqueness in the decomposition $H = M^\perp \oplus M^{\perp\perp}$ of Proposition 12.3, $m^\perp = 0$ and $m - u = 0$. Therefore $u = m$ is in M, and $M^{\perp\perp} = M$.

Theorem 12.5 (Riesz Representation Theorem). If ℓ is a continuous linear functional on the Hilbert space H, then there exists a unique v in H with $\ell(u) = (u, v)$ for all v in H. This vector v has the property that $\|\ell\| = \|v\|$.

REMARKS. It is instructive to compare this result with the version of the Riesz Representation Theorem in Theorem 9.19, which applies to $L^p(S, \mathcal{A}, \mu)$ for $1 \le p < \infty$ and in particular to $L^2(S, \mathcal{A}, \mu)$. That theorem associates to a continuous linear functional ℓ on this L^2 space a member g of the space such that $\ell(f) = \int_S fg\,d\mu$ for all f in the space. The present theorem, applied with $H = L^2(S, \mathcal{A}, \mu)$, instead yields a member v of the space such that $\ell(f) = \int_S f\bar{v}\,d\mu$

for all f in the space. The connection, of course, is that the function g is $\bar{v}$. The space $L^2(S, \mathcal{A}, \mu)$ has a canonically defined notion of complex conjugation, but an abstract Hilbert space does not. Because of the existence of this canonical conjugation, Theorem 9.19 gives us a canonical linear isometry of $L^2(S, \mathcal{A}, \mu)^*$ onto $L^2(S, \mathcal{A}, \mu)$, whereas Theorem 12.5 gives us a canonical isometry that is merely conjugate linear.

PROOF. Uniqueness is immediate since if $(u, v) = 0$ for all u, then $(u, v) = 0$ for $u = v$, and hence $v = 0$. Let us prove existence. If $\ell = 0$, take $v = 0$. Otherwise let $M = \{u \mid \ell(u) = 0\}$. This is a vector subspace since ℓ is linear, and it is closed since ℓ is continuous. By Proposition 12.3 and the fact that M is not all of H, $M^\perp$ contains a nonzero vector w. This vector w must have $\ell(w) \neq 0$ since $M \cap M^\perp = 0$, and we let v be the member of $M^\perp$ given by

$$v = \frac{\overline{\ell(w)}}{\|w\|^2}\, w.$$

For any u in H, we have $\ell\big(u - \frac{\ell(u)}{\ell(w)}\, w\big) = 0$, and hence $u - \frac{\ell(u)}{\ell(w)}\, w$ is in M. Since v is in $M^\perp$, $u - \frac{\ell(u)}{\ell(w)}\, w$ is orthogonal to v. Thus

$$(u, v) = \Big(\frac{\ell(u)}{\ell(w)}\, w,\ v\Big) = \Big(\frac{\ell(u)}{\ell(w)}\, w,\ \frac{\overline{\ell(w)}}{\|w\|^2}\, w\Big) = \ell(u)\,\frac{\ell(w)}{\ell(w)}\,\frac{\|w\|^2}{\|w\|^2} = \ell(u).$$

This proves existence.

For the norm equality every u in H has $|\ell(u)| = |(u, v)| \le \|u\|\|v\|$ by the Schwarz inequality. Taking the supremum over all u with $\|u\| \le 1$ gives $\|\ell\| \le \|v\|$. On the other hand, $|(u, v)| = |\ell(u)| \le \|\ell\|\|u\|$; putting $u = v$ gives $\|v\| \le \|\ell\|$. Thus $\|\ell\| = \|v\|$.

A subset S of H is **orthonormal** if each vector in S has norm 1 and if each pair of distinct vectors in S is orthogonal. For example, relative to the inner product $(f, g) = \frac{1}{2\pi}\int_\pi^\pi f\bar{g}\,dx$, the functions $x \mapsto e^{inx}$ are orthonormal as n varies through the integers. An orthonormal set S is linearly independent; in fact, if $v_1, \dots, v_n$ are members of S with $\sum_i c_i v_i = 0$, then the computation $0 = \big(v_j, \sum_i c_i v_i\big) = \sum_i c_i (v_j, v_i) = c_j\|v_j\|^2 = c_j$ shows that $c_j = 0$ for all j.

We encountered other examples of orthogonal sets, beyond the functions e^{inx}, in Chapter IV in connection with solving certain ordinary differential equations. Such an orthogonal set becomes orthonormal when each member is scaled by the reciprocal of its norm. One example was the system of Legendre polynomials $P_n(x)$, which were introduced in Section IV.8: the differential equation $(1-t^2)y'' - 2ty' + n(n+1)y = 0$ has polynomial solutions $y(t)$ that are unique up to a scalar, and $P_n(t)$ is a suitably normalized polynomial solution, necessarily of degree n. These can be shown to be orthogonal[3] in $L^2([-1,1], dt)$.

Another example was constructed from the Bessel function

$$J_0(t) = \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n}}{2^{2n}(n!)^2},$$

which was defined in Section IV.8. There are infinitely many distinct positive real numbers k_n such that $J_0(k_n) = 0$, and it can be shown that the functions $x \mapsto J_0(k_n x)$ are orthogonal[4] in $L^2([0,1],\ x\,dx)$.

If an ordered set of n linearly independent vectors in H is given, the **Gram–Schmidt orthogonalization process**, which appears in Problem 6 at the end of the present chapter, gives an algorithm for replacing the set with an orthonormal set having the same linear span.

Let M be a closed vector subspace of H, so that $H = M \oplus M^\perp$ by Proposition 12.3. The linear projection operator E of H on M along $M^\perp$, given by the identity on M and the 0 operator on $M^\perp$, is called the **orthogonal projection** of H on M. The linear operator E is bounded with $\|E\| \le 1$ because if $u \in H$ decomposes as $u = m + m^\perp$, the Pythagorean Theorem gives

$$\|E(u)\|^2 = \|E(m+m^\perp)\|^2 = \|m\|^2 \le \|m\|^2 + \|m^\perp\|^2 = \|u\|^2.$$

We are going to derive a formula for E in terms of orthonormal sets.

Lemma 12.6. If $\{u_j\}$ is an orthonormal sequence in the Hilbert space H and if $\{c_j\}$ is a sequence of scalars, then $\sum_{j=1}^\infty c_j u_j$ converges if and only if $\sum_{j=1}^\infty |c_j|^2 < \infty$, and in this case

$$\Big\| \sum_{j=1}^\infty c_j u_j \Big\| = \Big(\sum_{j=1}^\infty |c_j|^2 \Big)^{1/2}.$$

When the series converges, the sum $\sum_{j=1}^\infty c_j u_j$ is independent of the order of the terms.

PROOF. For $m \ge n$, we have

$$\Big\| \sum_{j=n}^m c_j u_j \Big\|^2 = \Big(\sum_{i=n}^m c_i u_i,\ \sum_{j=n}^m c_j u_j \Big) = \sum_{i,j} c_i \bar c_j (u_i, u_j) = \sum_{j=n}^m |c_j|^2.$$

This shows that the sequence $\Big\{ \sum_{j=1}^p c_j u_j \Big\}$ is Cauchy in H if and only if $\sum_{j=1}^\infty |c_j|^2$ is convergent, and the first conclusion follows since H is complete. When

[3]The verification appears in the problems in the companion volume *Advanced Real Analysis*.

[4]Again the verification appears in the problems in the companion volume *Advanced Real Analysis*.

$\left\{ \sum_{j=1}^p c_j u_j \right\}$ is convergent, we denote its limit by $\sum_{j=1}^\infty c_j u_j$, and continuity of the norm yields $\| \sum_{j=1}^\infty c_j u_j \| = \lim_p \| \sum_{j=1}^p c_j u_j \|$. Since we have seen that $\| \sum_{j=1}^p c_j u_j \| = \left(\sum_{j=1}^p |c_j|^2 \right)^{1/2}$, the second conclusion of the lemma follows.

Let $u = \sum_j c_j u_j$, and let $\sum_k c_{j_k} u_{j_k}$ be a rearrangement, necessarily convergent by what has already been proved. Suppose that the rearrangement has sum u'. The equality just proved shows that $\|u\|^2 = \sum_{i=1}^\infty |c_i|^2 = \|u'\|^2$ since rearrangements of series of nonnegative reals have the same sums. Continuity of the inner product, together with the same computation as made above, gives

$$(u, u') = \lim_{p,q} \left(\sum_{i=1}^p c_i u_i,\ \sum_{k=1}^q c_{j_k} u_{j_k} \right) = \lim_{p,q} \sum_{\substack{1 \le i \le p, \\ i = j_k \text{ with } k \le q}} |c_i|^2.$$

The limit on the right is $\sum_{i=1}^\infty |c_i|^2$ since $\sum_k |c_{j_k}|^2$ is a rearrangement of $\sum_i |c_i|^2$, and hence $(u, u') = \sum_{i=1}^\infty |c_i|^2 = \|u\|^2 = \|u'\|^2$. Therefore $\|u - u'\|^2 = (u, u) - 2\,\mathrm{Re}(u, u') + (u', u') = \|u\|^2 - 2\|u\|^2 + \|u\|^2 = 0$, and $u' = u$.

Proposition 12.7. Let S be an orthonormal set in the Hilbert space H, and let M be the smallest closed vector subspace of H containing S. For each u in H, there are at most countably many members v_α of S such that $(u, v_\alpha) \neq 0$, and thus the series

$$E(u) = \sum_{v_\alpha \in S} (u, v_\alpha) v_\alpha$$

has only countably many nonzero terms. The series converges independently of the order of the nonzero terms, E is the orthogonal projection of H on M, and E satisfies

$$\|E(u)\|^2 = \sum_{v_\alpha \in S} |(u, v_\alpha)|^2 \le \|u\|^2.$$

REMARK. The final inequality of the proposition is **Bessel's inequality**.

PROOF. Let $v_{\alpha_1}, \dots, v_{\alpha_n}$ be a finite subset of S, and form the vector $u' = \sum_{j=1}^n (u, v_{\alpha_j}) v_{\alpha_j}$. Taking the inner product of both sides with u gives

$$(u', u) = \sum_{j=1}^n (u, v_{\alpha_j})(v_{\alpha_j}, u) = \sum_{j=1}^n |(u, v_{\alpha_j})|^2,$$

and Lemma 12.6 gives

$$\|u'\|^2 = \sum_{j=1}^n |(u, v_{\alpha_j})|^2.$$

Therefore $0 \le \|u-u'\|^2 = \|u\|^2 - 2\,\mathrm{Re}(u, u') + \|u'\|^2 = \|u\|^2 - 2\|u'\|^2 + \|u'\|^2 = \|u\|^2 - \|u'\|^2$, and we obtain

$$\|u'\|^2 \le \|u\|^2. \tag{$*$}$$

In other words,

$$\sum_{j=1}^{n} |(u, v_{\alpha_j})|^2 \le \|u\|^2, \tag{$**$}$$

no matter what finite subset $v_{\alpha_1}, \dots, v_{\alpha_n}$ of S we use.

The sum of uncountably many positive real numbers is infinite, since otherwise there could be only finitely many greater than $1/n$ for each n. Since $\|u\|^2 < \infty$, $(**)$ implies that there can be only countably many α's with $|(u, v_\alpha)|^2$ nonzero. This proves the first conclusion. If we enumerate those α's and apply Lemma 12.6, we obtain the convergence of $\sum_{v_\alpha \in S}(u, v_\alpha)v_\alpha$ to a sum independent of the order of the terms.

It is evident from the formula that E is linear and that $E(u) = 0$ if u is in $M^\perp$. Inequality $(**)$ shows that the partial sums u' of $E(u)$ have $\|u'\| \le \|u\|$, and the continuity of the norm therefore implies that $\|E(u)\| \le \|u\|$ for all u. Hence E is continuous. Since $E(v_\alpha) = v_\alpha$ for all α, E is the identity on all finite linear combinations of members of S. The continuity of E thus implies that E is the identity on all of M. Hence E is the orthogonal projection as asserted. The final assertion of the proposition follows from Lemma 12.6 and the inequality $\|E(u)\| \le \|u\|$, which we have already proved.

Corollary 12.8. If S is an orthonormal set in the Hilbert space H, then the following are equivalent:

(a) S is maximal among orthonormal subsets of H,

(b) $u = \sum_{v_\alpha \in S} (u, v_\alpha) v_\alpha \quad$ for all u in H,

(c) $\|u\|^2 = \sum_{v_\alpha \in S} |(u, v_\alpha)|^2 \quad$ for all u in H,

(d) $(u, v) = \sum_{v_\alpha \in S} (u, v_\alpha)\overline{(v, v_\alpha)} \quad$ for all u and v in H.

REMARKS. Condition (b) is summarized by saying that the orthonormal set S is an **orthonormal basis** of H. If H is infinite-dimensional, an orthonormal basis is not a basis in the ordinary linear-algebra sense; a passage to the limit is usually needed to expand vectors in terms of the basis. Condition (c), or sometimes condition (d), is called **Parseval's equality**. Thus the corollary says that the orthonormal set S is maximal if and only if it is an orthonormal basis, if and only if Parseval's equality holds.

PROOF. Let M be the smallest closed vector subspace of H containing S. Then S is maximal if and only if $M^\perp = 0$, and we replace (a) by this condition. If $M^\perp = 0$, then E is the identity operator in Proposition 12.7, and the proposition shows that (b) holds. If (b) holds, Proposition 12.7 says that (c) holds. On the other hand, if (c) holds, then Proposition 12.7 says that $\|u\| = \|E(u)\|$ for all u. For a vector u in $M^\perp$, which must have $E(u) = 0$, this says that $\|u\| = 0$. Thus $M^\perp = 0$, and (a) holds. Hence (a), (b), and (c) are equivalent. Finally (c) and (d) are equivalent by polarization.

In the context of Fourier series, Parseval's equality ((c) in Corollary 12.8) was proved as Theorem 6.49, and that theorem showed also that any member of $L^2\big([-\pi, \pi],\ \frac{1}{2\pi}\,dx\big)$ is the sum of its Fourier series in the sense of convergence in L^2. This conclusion was (b) in the corollary. The corollary is showing that the equivalence of (b) and (c) is just a result in abstract Hilbert-space theory. The extra content of Theorem 6.49 is that these conditions are actually satisfied by the system of exponential functions.

One can show that the other two examples we gave in this section of orthogonal sets give orthonormal bases when normalized—the Legendre polynomials $P_n(t)$ on $[-1, 1]$ with respect to dt and the functions $J_0(k_n t)$ on $[0, 1]$ with respect to $t\,dt$.

Proposition 12.9. Let (X, μ) and (Y, ν) be σ-finite measure spaces, and suppose that $L^2(X, \mu)$ has a countable orthonormal basis $\{u_i\}$ and $L^2(Y, \nu)$ has a countable orthonormal basis $\{v_j\}$. Then $\{(x, y) \mapsto u_i(x)v_j(y)\}$ is an orthonormal basis of $L^2(X \times Y, \mu \times \nu)$.

PROOF. The functions $u_i(x)v_j(y)$ are orthonormal, and Corollary 12.8 shows that it is enough to prove that this orthonormal set is maximal. Suppose that $w(x, y)$ is an L^2 function on $X \times Y$ orthogonal to all of them. Then

$$0 = \int_X \int_Y w(x, y)\,\overline{u_i(x)}\,\overline{v_j(y)}\,d\nu(y)\,d\mu(x) = \int_X (w(x,\ \cdot\,), v_j)\,\overline{u_i(x)}\,d\mu(x)$$

for all i and j. Since $\{u_i\}$ is an orthonormal basis of $L^2(X, \mu)$, $x \mapsto (w(x,\ \cdot\,), v_j)$ is the 0 function in $L^2(X, \mu)$ for each j. In other words, $(w(x,\ \cdot\,), v_j) = 0$ for a.e. x $[d\mu]$ for that j. Since the number of j's is countable, $(w(x,\ \cdot\,), v_j) = 0$ for all j for a.e. x $[d\mu]$. Any such x has $0 = \sum_j |(w(x,\ \cdot\,), v_j)|^2 = \int_Y |w(x, y)|^2\,d\nu(y)$. Integrating in x, we see that w is the 0 function in $L^2(X \times Y, \mu \times \nu)$.

Proposition 12.10. Any orthonormal set in a closed vector subspace M of a Hilbert space H can be extended to an orthonormal basis of M. In particular any closed vector subspace M of H has an orthonormal basis.

PROOF. As a closed subset of a complete space, M is complete, and therefore M is a Hilbert space in its own right. Order by inclusion all orthonormal subsets of M containing the given set. The given set is one such, and the union of the members of a chain is an orthonormal set forming an upper bound for the chain. By Zorn's Lemma we can find a maximal orthonormal set S in M containing the given one. This satisfies (a) in Corollary 12.8 and hence is an orthonormal basis. This proves the first conclusion, and the second conclusion follows from the first by taking the given orthonormal set in M to be empty.

Proposition 12.11. Any two orthonormal bases of a Hilbert space have the same cardinality.

REMARKS. Cardinality is discussed in Section A10 of the appendix. The "same cardinality" whose existence is proved in the proposition is called the **Hilbert space dimension** of the Hilbert space. Problem 7 at the end of the chapter shows that two Hilbert spaces are isomorphic as Hilbert spaces if and only if they have the same Hilbert space dimension. Despite the apparent definitive sound of this result, one must not attach too much significance to the proposition. Hilbert spaces that arise in practice tend to have some additional structure, and an isomorphism of this kind need not preserve the additional structure.

PROOF. Fix two orthonormal bases $U = \{u_\alpha\}$ and $V = \{v_\beta\}$ of a Hilbert space H. We define two members u_α and $u_{\alpha'}$ of U to be equivalent if there exists a sequence

$$u_{\alpha_1}, v_{\beta_1}, u_{\alpha_2}, v_{\beta_2}, \dots, u_{\alpha_{n-1}}, v_{\beta_{n-1}}, u_{\alpha_n} \qquad (*)$$

with $u_{\alpha_1} = u_\alpha$ and $u_{\alpha_n} = u_{\alpha'}$, with each u_{α_j} in U and each v_{β_j} in V, and with each consecutive pair having nonzero inner product. Define an equivalence relation in V similarly.

Each equivalence class is countable. In fact, consider the class of u_{α_1}, and consider sequences of a fixed length. Proposition 12.7 shows that only countably many members of V can have nonzero inner product with u_{α_1}, only countably many members of U can have nonzero inner product with that, and so on. Thus there are only countably many sequences of any particular length. The countable union of these countable sets is countable, and thus there are only finitely many sequences connecting u_{α_1} to anything. Hence u_{α_1} can be equivalent to only countably many members of U.

Let U_1 and V_1 be equivalence classes in U and V, respectively, and suppose that u_{α_0} and v_{β_0} are members of U_1 and V_1 with nonzero inner product. Expand u_{α_0} in terms of V as $u_{\alpha_0} = \sum_\beta (u_{\alpha_0}, v_\beta)v_\beta$, retaining only the terms with $(u_{\alpha_0}, v_\beta) \neq 0$. One of the terms making a contribution is the one with $v_\beta = v_{\beta_0}$, and it follows that any other term with $(u_{\alpha_0}, v_{\beta'}) \neq 0$ has $v_{\beta'}$ equivalent to v_β. Hence we have

$$u_{\alpha_0} = \sum_{v_\beta \in V_1} (u_{\alpha_0}, v_\beta) v_\beta \qquad \text{and similarly} \qquad v_{\beta_0} = \sum_{u_\alpha \in U_1} (v_{\beta_0}, u_\alpha) u_\alpha.$$

If $u_{\alpha_0'}$ is another member of U_1 and we expand it in terms of V, retaining only the nonzero terms, then the v_β's that occur have to be equivalent to one another. So we have $u_{\alpha_0'} = \sum_{v_\beta \in V_2} (u_{\alpha_0}, v_\beta) v_\beta$ for some equivalence class V_2 within V. If we form a sequence $(*)$ connecting u_{α_0} and $u_{\alpha_0'}$, we see that at least one member of V_2 is connected to at least one member of V_1. Thus $V_1 = V_2$. Consequently every member of U_1 lies in the smallest closed vector subspace containing V_1, and every member of V_1 lies in the smallest closed subspace containing U_1. In other words, U_1 and V_1 are orthonormal bases for the same closed vector subspace of H.

If U_1 is finite, then linear algebra shows that V_1 is finite and has the same number of elements. Since U_1 and V_1 are countable, the only way that either can be infinite is if both are countably infinite. In any event, U_1 and V_1 have the same cardinality. Thus we have a one-one function carrying U_1 onto V_1. Repeating this process for each equivalence class within U, we obtain a one-one function carrying U onto V.

3. Bounded Linear Operators on Hilbert Spaces

In this section we briefly study bounded linear operators from a Hilbert space H to itself. In the finite-dimensional case we often make a correspondence between matrices and linear operators by using the standard basis of the space of column vectors. If $\{e_i\}_{i=1}^n$ is this basis, then the correspondence between a matrix $A = [A_{ij}]$ and a linear operator L is given by $A_{ij} = (L(e_j), e_i)$. If $u = \sum_j u_j e_j$ and $v = \sum_i v_i e_i$ are column vectors, then $L(u) = \sum_j u_j L(e_j)$ and hence $(L(u), v) = \sum_{i,j} u_j \bar{v}_i (L(e_j), e_i) = \sum_{ij} \bar{v}_i A_{ij} u_j$.

We could extend these formulas to the case of a general Hilbert space, not necessarily finite-dimensional, by using a particular orthonormal basis as the generalization of $\{e_i\}$. But no particular such basis recommends itself, and we work without any choice of basis as much as possible, except for purposes of motivation. Instead, we may think of the function $(u, v) \mapsto (L(u), v)$ as a more appropriate—and canonical—analog of the matrix of L. Just as the operator norm of L is given by a formula that views L as an operator, namely

$$\|L\| = \sup_{\|u\| \leq 1} \|L(u)\|,$$

so there is a formula for computing the norm in terms of the function of two variables, namely

$$\|L\| = \sup_{\substack{\|u\|\le 1,\\ \|v\|\le 1}} |(L(u), v)|.$$

To verify this formula, fix u and let v have norm ≤ 1. Application of the Schwarz inequality gives $|(L(u), v)| \le \|L(u)\|\|v\| \le \|L(u)\|$. On the other hand, if $L(u) \ne 0$, we take $v = \|L(u)\|^{-1}L(u)$; this v has $\|v\| = 1$, and we obtain $|(L(u), v)| = \|L(u)\|^{-1}(L(u), L(u)) = \|L(u)\|$. Hence $\sup_{\|v\|\le 1} |(L(u), v)\| = \|L(u)\|$. Taking the supremum over $\|u\| \le 1$ shows that the two expressions for $\|L\|$ are equal.

We shall work with the "adjoint" L^* of a bounded linear operator L. In terms of matrices in the finite-dimensional case, the matrix of L^* is to be the conjugate transpose of the matrix of L. In other words, the $(i, j)^{\text{th}}$ entry $(L^*(e_j), e_i))$ of the matrix for L^* is to be $\overline{(L(e_i), e_j)} = (e_j, L(e_i))$. Passing to our functions of two variables, we want to arrange that $(L^*(u), v) = (u, L(v))$ for all u and v. Let us prove existence and uniqueness of such a bounded linear operator.

Proposition 12.12. Let $L : H \to H$ be a bounded linear operator on the Hilbert space H. For each u in H, there exists a unique vector $L^*(u)$ in H such that

$$(L^*(u), v) = (u, L(v)) \qquad \text{for all } v \text{ in } H.$$

As u varies, this formula defines L^* as a bounded linear operator on H, and $\|L^*\| = \|L\|$.

PROOF. The function $v \mapsto (L(v), u)$ is a linear functional on H satisfying $|(L(v), u)| \le \|L\|\|u\|\|v\|$, hence having norm $\le \|L\|\|u\|$. Being bounded, the linear functional is given by $(L(v), u) = (v, w)$ for some unique w in H, according to Theorem 12.5. We define $L^*(u) = w$, and then we have $(L^*(u), v) = (u, L(v))$. This formula shows that L^* is a linear operator, and the computation

$$\|L^*\| = \sup_{\substack{\|u\|\le 1,\\ \|v\|\le 1}} |(L^*(u), v)| = \sup_{\substack{\|u\|\le 1,\\ \|v\|\le 1}} |(u, L(v))| = \sup_{\substack{\|u\|\le 1,\\ \|v\|\le 1}} |(L(v), u)| = \|L\|$$

shows that $\|L^*\| = \|L\|$.

The bounded linear operator L^* in the proposition is called the **adjoint** of L. The mapping $L \mapsto L^*$ is conjugate linear. We shall be especially interested in the case that $L^* = L$, in which case we say that L is **self adjoint**.

An example of a self-adjoint operator is the orthogonal projection E on a closed vector subspace M as defined before Lemma 3.6. In fact, if u in H decomposes according to $H = M \oplus M^\perp$ as $u = u' + u''$, then the computation $(1 - E)(u) =$

$u - u' = u''$ shows that $1 - E$ is the orthogonal projection on $M^\perp$. Hence $(E(u), (1 - E)(v)) = 0$ for all u and v in H, and also $((1 - E)(u), E(v)) = 0$. The first of these says that $(E(u), v) = (E(u), E(v))$, and the second says that $(E(u), E(v)) = (u, E(v))$. Combining these, we obtain $(E(u), v) = (u, E(v))$. Comparison of this formula with the formula in Proposition 12.12 shows that $E = E^*$.

The Banach space $\mathcal{B}(H, H)$ is closed under composition. In fact, if L and M are in $\mathcal{B}(H, H)$, then linear algebra shows LM to be linear, and the computation $\|(LM)(u)\| = \|L(M(u))\| \leq \|L\|\|M(u)\| \leq \|L\|\|M\|\|u\|$ shows that

$$\|LM\| \leq \|L\|\|M\|.$$

Hence LM is in $\mathcal{B}(H, H)$ if L and M are. Within $\mathcal{B}(H, H)$, we have $(LM)^* = M^*L^*$.

4. Hahn–Banach Theorem

We return now to the setting of general normed linear spaces or Banach spaces. There are three main theorems concerning the norm topology of such spaces—the Hahn–Banach Theorem, the Uniform Boundedness Theorem, and the Interior Mapping Principle. These three theorems are the main subject matter of the remainder of this chapter.

We shall often use symbols $x, y, \ldots$ for members of a normed linear space and symbols $x^*, y^*, \ldots$ for linear functionals. This notation has the advantage of allowing us to use symbols like x^{**} for linear functionals on a space of linear functionals, an important notion as we shall see.

We begin with the Hahn–Banach Theorem, which ensures the existence of many continuous linear functionals on a normed linear space. The theorem has applications even in situations in which one has a concrete realization of the dual space, because it shows that any closed vector subspace is characterized by the continuous linear functionals that vanish on the subspace.

Theorem 12.13 (Hahn–Banach Theorem). If Y is a vector subspace of a normed linear space X and if y^* is a continuous linear functional on Y, then there exists a continuous linear functional x^* on X with $\|x^*\| = \|y^*\|$ such that

$$x^*(y) = y^*(y) \qquad \text{for all } y \in Y.$$

The theorem as stated is derived from the following lemma, which itself goes under the name "Hahn–Banach Theorem" and has other applications quite distinct from Theorem 12.13 that are beyond the scope of this book.

Lemma 12.14. Let X be a real vector space, and let p be a real-valued function on X with

$$p(x+x') \le p(x) + p(x') \qquad \text{and} \qquad p(tx) = tp(x)$$

for all x and x' in X and all real $t \ge 0$. If f is a linear functional on a vector subspace Y of X with $f(y) \le p(y)$ for all y in Y, then there exists a linear functional F on X with $F(y) = f(y)$ for all $y \in Y$ and $F(x) \le p(x)$ for all $x \in X$.

PROOF. Form the collection of all linear functionals on vector subspaces of X that extend f and that are dominated by p, and partially order the collection by saying that one is $\le$ another if the second is an extension of the first. If we have a chain of such extensions, then we can obtain an upper bound for the chain by taking the union of the domains and using the common value of the linear functionals on an element of this domain as the value of the linear functional forming the upper bound. The result is linear because any two members of the domain must lie in the domain of a single member of the chain. By Zorn's Lemma let f_0, with domain Y_0, be a maximal extension. We shall prove that $Y_0 = X$.

In fact, suppose that y_1 is a vector in X but not Y_0. Every vector in the vector subspace Y_1 spanned by y_1 and Y_0 has a unique representation as $y + cy_1$, where y is in Y_0 and c is in $\mathbb{R}$. Define f_1 on Y_1 by

$$f_1(y + cy_1) = f_0(y) + ck, \tag{$*$}$$

where k is a real number to be specified. For a suitable choice of k, f_1 will be bounded by p and will contradict the maximality of (f_0, Y_0).

Let y and y' be in Y_0. Then

$$f_0(y') - f_0(y) = f_0(y'-y) \le p(y'-y) \le p(y'+y_1) + p(-y_1-y),$$

and hence $$-p(-y_1-y) - f_0(y) \le p(y'+y_1) - f_0(y').$$

Take the supremum of the left side over y and the infimum of the right side over y', let k be any real number in between, and define f_1 on Y_1 by $(*)$.

To complete the proof, we are to check that $f_1(x) \le p(x)$ for all x in Y_1. Thus suppose that $x = y + cy_1$ is arbitrary in Y_1. If $c = 0$, then $f_1(x) \le p(x)$ by the assumption on Y_0. If $c > 0$, then

$$f_1(x) = f_0(y)+ck \le f_0(y)+c[p(c^{-1}y+y_1)-f_0(c^{-1}y)] = p(y+cy_1) = p(x).$$

If $c < 0$, then

$$f_1(x)=f_0(y)+ck\le f_0(y)+c[-p(-y_1-c^{-1}y)-f_0(c^{-1}y)]=p(y+cy_1)=p(x).$$

In any case, $f_1(x) \le p(x)$.

PROOF OF THEOREM 12.13. If the field of scalars is $\mathbb{R}$, then Theorem 12.13 follows immediately from Lemma 12.14 with $p(x) = \|y^*\|\|x\|$ and $f = y^*$.

If the field of scalars is $\mathbb{C}$, if y^* is given, and if, as we may, we regard X as a real normed linear space, then $\operatorname{Re} y^*$ defined by $(\operatorname{Re} y^*)(y) = \operatorname{Re}(y^*(y))$ is a real linear functional on Y with

$$|(\operatorname{Re} y^*)(y)| \le |y^*(y)| \le \|y^*\|\|y\| \qquad \text{for all } y \in Y.$$

By what has already been proved, we can extend $\operatorname{Re} y^*$ without an increase in norm to a real linear functional F defined on all of X. Define

$$x^*(x) = F(x) - iF(ix).$$

We show that x^* has the required properties. Certainly $x^*(x+x') = x^*(x)+x^*(x')$ and $x^*(cx) = cx^*(x)$ for c real. Furthermore

$$x^*(ix) = F(ix) - iF(i^2x) = i[F(x) - iF(ix)] = ix^*(x).$$

Thus x^* is complex linear. On Y, we have

$$(\operatorname{Re} y^*)(iy) + i(\operatorname{Im} y^*)(iy) = y^*(iy) = iy^*(y) = -(\operatorname{Im} y^*)(y) + i(\operatorname{Re} y^*)(y),$$

and thus $(\operatorname{Re} y^*)(iy) = -(\operatorname{Im} y^*)(y)$. Substituting this identity into the definition of x^*, we obtain

$$x^*(y) = (\operatorname{Re} y^*)(y) - i(\operatorname{Re} y^*)(iy) = (\operatorname{Re} y^*)(y) + i(\operatorname{Im} y^*)(y) = y^*(y)$$

for y in Y. Thus x^* is an extension of y^*. Finally if $x^*(x) = re^{i\theta}$ for r and θ real and $r \ge 0$, then

$$|x^*(x)| = x^*(e^{-i\theta}x) = F(e^{-i\theta}x) \le \|y^*\|\|e^{-i\theta}x\| = \|y^*\|\|x\|,$$

since the nonnegative number $x^*(e^{-i\theta}x)$ has 0 imaginary part. Thus $\|x^*\| \le \|y^*\|$. The reverse inequality follows because x^* is an extension of y^*, and the proof is complete.

Corollary 12.15. If Y is a closed vector subspace of a normed linear space X and if x_0 is a vector of X not in Y, then there exists an x^* in the dual X^* with

$$x^*(y) = 0 \qquad \text{for all } y \in Y$$

and

$$x^*(x_0) = 1.$$

The norm of x^* can be taken to be the reciprocal of the distance from x_0 to Y.

PROOF. Let $d > 0$ be the distance from x_0 to Y, and let Z be the linear span of x_0 and Y. Every x in Z has a unique expansion as $x = y + cx_0$ for some scalar c and some y in Y. For such an x, let $z^*(x) = c$. Let us see that the linear function z^* on Z satisfies

$$\|z^*\| = d^{-1}. \tag{$*$}$$

First we check that $|z^*(x)| \le d^{-1}\|x\|$: if $c \neq 0$, then

$$\|x\| = \|y + cx_0\| = |c|\|c^{-1}y + x_0\| \ge |c|d = d|z^*(x)|,$$

while if $c = 0$, then $z^*(x) = 0$. Thus $|z^*(x)| \le d^{-1}\|x\|$ for all x, and we obtain $\|z^*\| \le d^{-1}$. For the reverse inequality, let $\{y_n\}$ be a sequence in Y, not necessarily convergent, with $\lim_n \|x_0 - y_n\| = d$. Then

$$1 = z^*(x_0 - y_n) \le \|z^*\|\|x_0 - y_n\| \longrightarrow d\|z^*\|,$$

and hence $\|z^*\| \ge d^{-1}$. This proves $(*)$. Applying Theorem 12.13 to z^*, we obtain the corollary.

EXAMPLE. To illustrate Corollary 12.15, we re-prove the result of Proposition 11.21a that $C(S)$ is dense in $L^p(S, \mu)$ if S is a compact Hausdorff space, μ is a regular Borel measure on S, and p satisfies $1 \le p < \infty$. For definiteness let us suppose that the underlying scalars are real. If $C(S)$ were not dense, then the corollary would produce a continuous linear functional ℓ on $L^p(S, \mu)$ that vanishes on $C(S)$ but is not identically 0 on $L^p(S, \mu)$. Theorem 9.19 says that ℓ has to be given by integration with some member g of $L^{p'}(S, \mu)$, where p' is the dual index: $\ell(f) = \int_S fg\,d\mu$ for all f in $L^p(S, \mu)$. Since ℓ vanishes on $C(S)$, we have $\int_S fg\,d\mu = 0$ for all $f \in C(S)$. Thus $\int_S fg^+\,d\mu = \int_S fg^-\,d\mu$ for all $f \in C(S)$. Here $g^+\,d\mu$ and $g^-\,d\mu$ are Borel measures on S, regular by Proposition 11.20, and they yield the same positive linear functional on $C(S)$. Applying the uniqueness in the Riesz Representation Theorem (Theorem 11.1), we obtain $g^+\,d\mu = g^-\,d\mu$ and therefore $g^+ = g^-$ almost everywhere. Since g^+ and g^- are nowhere both nonzero, $g^+ = g^- = 0$ almost everywhere. Hence g is the 0 function, and $\ell = 0$, contradiction.

Corollary 12.16. If X is a normed linear space and if $x_0 \neq 0$ is a vector in X, then there is an x^* in X^* with

$$\|x^*\| = 1 \qquad \text{and} \qquad x^*(x_0) = \|x_0\|.$$

PROOF. Apply Corollary 12.15 with $Y = 0$ and multiply by $\|x_0\|$ the linear functional that is produced by that corollary.

Corollary 12.16, when applied to $x_0 = x - x'$, shows that there are enough continuous linear functionals on a normed linear space X to separate points. Also, it implies that the only vector x_0 in X with $x^*(x_0) = 0$ for all x^* in X^* is $x_0 = 0$. The third corollary we have already seen for L^p spaces with $1 \le p < \infty$ in Proposition 9.8, at least when the measure space is σ-finite.

Corollary 12.17. If X is a normed linear space and x_0 is in X, then

$$\|x_0\| = \sup_{\|x^*\|\le 1} |x^*(x_0)|.$$

PROOF. If $\|x^*\| \le 1$, then $|x^*(x_0)| \le \|x^*\|\|x_0\| \le \|x_0\|$, and therefore $\sup_{\|x^*\|\le 1} |x^*(x_0)| \le \|x_0\|$. The linear functional of Corollary 12.16 shows that equality holds.

We have seen for σ-finite measure spaces that $X = L^1(S, \mu)$ may be identified with $L^\infty(S, \mu)$ via integration. In turn every member of $L^1(S, \mu)$ then acts as a continuous linear functional on $L^\infty(S, \mu)$ via integration. This change of point of view amounts to the implementation of a certain canonically defined linear mapping of X into X^{**}, which we now define for general normed linear spaces.

Let X be a normed linear space, and let X^{**} be the dual of X^*. We define a linear operator $\iota : X \to X^{**}$ by

$$(\iota(x))(x^*) = x^*(x) \qquad \text{for all } x^* \in X^*,$$

and we call ι the **canonical map** of X into X^{**}.

Corollary 12.18. If X is a normed linear space, then the canonical map $\iota : X \to X^{**}$ has $\|\iota(x)\| = \|x\|$ for all x and in particular is one-one. Consequently if X is complete, then $\iota(X)$ is a closed vector subspace of X^{**}.

PROOF. We have

$$\|\iota(x)\| = \sup_{\|x^*\|\le 1} |(\iota(x))(x^*)| = \sup_{\|x^*\|\le 1} |x^*(x)| = \|x\|,$$

the last step holding by Corollary 12.17. This proves the first conclusion. Because ι preserves norms, X complete implies that $\iota(X)$ is a complete subset of the complete space X^{**} and is therefore closed, by Corollary 2.43.

A Banach space X is said to be **reflexive** if the canonical map carries X onto X^{**}. *Warning:* This is a more restrictive condition than to say that there is some norm-preserving linear mapping of X onto X^{**}.

Finite-dimensional normed linear spaces are reflexive since linear functionals in this case are automatically continuous and since the vector-space dual of a finite-dimensional vector space has the same dimension as the space itself. Hilbert spaces are reflexive as a consequence of the Riesz Representation Theorem in its form in Theorem 12.5. The spaces $L^p(S, \mu)$ for a σ-finite measure space, when $1 < p < \infty$, are reflexive as a consequence of the Riesz Representation Theorem[5] in its form in Theorem 9.19. However, $L^1(S, \mu)$ and $L^\infty(S, \mu)$ are often not reflexive, as is shown below in Proposition 12.19 and Corollary 12.21.

Proposition 12.19. If (S, μ) is a σ-finite measure space with infinitely many disjoint sets of positive measure, then $L^1(S, \mu)$ is not reflexive.

PROOF. Theorem 9.19 shows that the Banach space $X = L^1(S, \mu)$ has $X^* \cong L^\infty(S, \mu)$, the isomorphism being given by integration. Therefore it is enough to produce a continuous linear functional on $L^\infty(S, \mu)$ that is not given by integration with an L^1 function.

Thus let $\{E_n\}$ be a sequence of disjoint sets of positive measure, and let Y be the vector subspace of functions in $L^\infty(S, \mu)$ that are constant on each E_n and have values on the E_n's tending to a finite limit as n tends to infinity. Let y^* of such a function be the limit. Then y^* is a linear functional on Y of norm 1. By the Hahn–Banach Theorem (Theorem 12.13), there exists a linear functional x^* defined on all of $L^\infty(S, \mu)$, having norm 1, and restricting to y^* on Y. Suppose that there is some g in $L^1(S, \mu)$ with $x^*(f) = \int_S fg\,d\mu$ for all f in Y, quite apart from all f in $L^\infty(S, \mu)$. If f is 1 on E_n and is 0 elsewhere, then $x^*(f) = 0$, and hence $\int_{E_n} g\,d\mu = 0$. In other words, $\int_{E_n} g\,d\mu = 0$ for every n. If we next take f to be 1 on $\bigcup_{n=1}^\infty E_n$ and to be 0 elsewhere, then $x^*(f) = 1$. On the other hand, this f has

$$x^*(f) = \int_S fg\,d\mu = \int_{\bigcup_n E_n} g\,d\mu = \sum_{n=1}^\infty \int_{E_n} g\,d\mu = 0,$$

and we have a contradiction.

Proposition 12.20. If X is a Banach space and its dual X^* is reflexive, then X is reflexive.

PROOF. Let $\iota : X \to X^{**}$ and $\iota^* : X^* \to X^{***}$ be the canonical maps. Arguing by contradiction, suppose that X is not reflexive. Since $\iota(X)$ is a closed proper vector subspace of X^{**}, Corollary 12.15 produces a nonzero member x^{***} of X^{***} such that $x^{***}(\iota(X)) = 0$. Since X^* is reflexive by assumption, there exists x^* in X^* with $x^{***} = \iota^*(x^*)$. If x is in X, then we have $0 = x^{***}(\iota(x)) = (\iota^*(x^*))(\iota(x)) = (\iota(x))(x^*) = x^*(x)$, and hence $x^* = 0$. But then $x^{***} = \iota^*(x^*) = 0$, and we have a contradiction.

[5] Actually, the σ-finiteness is not needed for $1 < p < \infty$.

Corollary 12.21. If (S, μ) is a σ-finite measure space with infinitely many disjoint sets of positive measure, then $L^\infty(S, \mu)$ is not reflexive.

PROOF. Theorem 9.19 shows that the Banach space $X = L^1(S, \mu)$ has $X^* \cong L^\infty(S, \mu)$, the isomorphism being given by integration. If X^* were reflexive, then X would have to be reflexive by Proposition 12.20, in contradiction to Proposition 12.19.

5. Uniform Boundedness Theorem

The second main theorem about the norm topology of normed linear spaces is the Uniform Boundedness Theorem, also known as the Banach–Steinhaus Theorem. This result involves a parametrized family of linear operators from one normed linear space into another, and it is assumed that the domain is complete. Two kinds of boundedness as a function of one variable are assumed—boundedness of each linear operator as a function on (the unit ball of) the domain and boundedness in the parameter for each fixed member of the domain. The conclusion is boundedness in the two variables jointly.

Theorem 12.22 (Uniform Boundedness Theorem). If $\{L_\alpha\}$ is a set of bounded linear operators from a Banach space X into a normed linear space Y such that

$$\|L_\alpha(x)\| \le C_x \qquad \text{for all } \alpha,$$

then there is a constant C independent of x such that $\|L_\alpha\| \le C$ for all α.

PROOF. For each positive integer n, the set

$$F_n = \big\{x \in X \;\big|\; \|L_\alpha(x)\| \le n \text{ for all } \alpha\big\}$$

is closed in X, being the intersection of inverse images of closed sets in Y under continuous functions, and $\bigcup_{n=1}^\infty F_n = X$ by assumption. By the Baire Category Theorem (Theorem 2.53b), one of the sets, say F_N, contains a nonempty open subset B of X. Then $\|L_\alpha(x)\| \le N$ for all α and for all x in B. If B contains the open ball in X of radius $2r > 0$ and center b, then $\|x\| \le r$ implies that $x + b$ is in B and that

$$\|L_\alpha(x)\| = \|L_\alpha(x+b) - L_\alpha(b)\| \le \|L_\alpha(x+b)\| + \|L_\alpha(b)\| \le N + C_b,$$

independently of α. Hence $\|x\| \le 1$ implies

$$\|L_\alpha(x)\| = r^{-1}\|L_\alpha(rx)\| \le r^{-1}(N + C_b).$$

In other words, $\|L_\alpha\| \le r^{-1}(N + C_b)$.

EXAMPLE. Let us use the theorem to give a proof that the Fourier series of a continuous periodic function need not converge at some point. Consider the Banach space X of all continuous periodic functions f on $[-\pi, \pi]$ with the supremum norm. Let D_n be the Dirichlet kernel as in Section I.10, given by

$$D_n(t) = \sum_{k=-n}^{n} e^{ikt} = \frac{\sin((n+\frac{1}{2})t)}{\sin\frac{1}{2}t}.$$

The n^{th} partial sum of the Fourier series of f is

$$s_n(f; x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_n(t)\,dt.$$

Define linear functionals ℓ_n on X by

$$\ell_n(f) = s_n(f; 0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(-t)D_n(t)\,dt.$$

Each of these is bounded; specifically $\|\ell_n\| \le 2n+1$ because $\|D_n\|_{\sup} \le 2n+1$. If the Fourier series of each continuous function f were to converge at 0, then $\lim_n \ell_n(f)$ would exist for each f, and hence we would have $|\ell_n(f)| \le C_f$ for a constant C_f independent of n. The Uniform Boundedness Theorem would say that $\|\ell_n\| \le C$ for some constant C independent of n. The norm equality of Theorem 11.26 or 11.28 would then allow us to conclude that $\int_{-\pi}^{\pi} |D_n(t)|\,dt$ is bounded. In fact, the numbers $\int_{-\pi}^{\pi} |D_n(t)|\,dt$ are unbounded, according to the following proposition, and thus there exists a continuous periodic function whose Fourier series diverges at $x = 0$.

Proposition 12.23. The numbers

$$L_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\,dt$$

have the property that

$$L_n = 4\pi^{-2}\log n + O(1),$$

where $O(1)$ denotes an expression bounded as a function of n. Hence L_n is unbounded with n.

REMARK. The numbers L_n are sometimes called **Lebesgue constants**.

PROOF. By writing $\sin((n+\frac{1}{2})t) = \sin nt \cos\frac{1}{2}t + \cos nt \sin\frac{1}{2}t$, we see that

$$D_n(t) = \sin nt \cot\tfrac{1}{2}t + \cos nt = 2t^{-1}\sin nt + h_n(t),$$

where $h_n(t)$ is bounded in the pair (n,t) for $|t| \le \pi$. If we let $O(1)$ denote an expression bounded as a function of n, then

$$\begin{aligned}
L_n &= \frac{2}{2\pi}\int_{-\pi}^{\pi}\frac{|\sin nt|}{|t|}\,dt + O(1)\\
&= \frac{2}{\pi}\int_0^{\pi}\frac{|\sin nt|}{t}\,dt + O(1)\\
&= \frac{2}{\pi}\sum_{k=0}^{n-1}\int_{k\pi/n}^{(k+1)\pi/n}\frac{|\sin nt|}{t}\,dt + O(1)\\
&= \frac{2}{\pi}\int_0^{\pi/n}\frac{\sin nt}{t}\,dt + \frac{2}{\pi}\int_0^{\pi/n}(\sin nt)\Big[\sum_{k=1}^{n-1}\frac{1}{t+k\pi/n}\Big]\,dt + O(1).
\end{aligned}$$

The first term on the right side is bounded, and the sum in brackets lies between

$$\pi^{-1}n(1+\tfrac{1}{2}+\cdots+\tfrac{1}{n-1}) \qquad \text{and} \qquad \pi^{-1}n(\tfrac{1}{2}+\cdots+\tfrac{1}{n}),$$

which are upper and lower Riemann sums for $\pi^{-1}n\int_1^n t^{-1}\,dt$ and have difference $\pi^{-1}n(1-\frac{1}{n})$. Thus the sum in brackets is equal to $\pi^{-1}n(\log n + O(1))$. The integral of $\sin nt$ over $[0,\pi/n]$ is $2/n$, and the result follows.

6. Interior Mapping Principle

The third main theorem about the norm topology of normed linear spaces is the Interior Mapping Principle. This result involves a single bounded linear operator from one normed linear space into another, and it is assumed that the domain and the range are both complete. The theorem is that if the operator is onto the range, then it carries open sets to open sets.

Theorem 12.24 (Interior Mapping Principle). If L is a continuous linear operator from a Banach space X *onto* a Banach space Y, then L carries open subsets of X to open subsets of Y.

PROOF. Let B_r be the closed ball in X with center 0 and radius r, and let U_s be the open ball in Y with center 0 and radius s. The proof is in three steps.

The first step is to show that $(L(B_1))^{\mathrm{cl}}$ contains an open neighborhood of 0 in Y. To do so, we use the fact that L is onto Y to write

$$Y = L(X) = L\big(\bigcup\nolimits_{n=1}^{\infty} B_n\big) = \bigcup\nolimits_{n=1}^{\infty} L(B_n).$$

Thus $Y = \bigcup_{n=1}^{\infty}(L(B_n))^{\mathrm{cl}}$, and the Baire Category Theorem (Theorem 2.53b) shows that one of the sets $(L(B_n))^{\mathrm{cl}}$ contains a nonempty open set. Since L is linear and since multiplication by $2n$ is a homeomorphism of Y, $(L(B_n))^{\mathrm{cl}} = (L(2nB_{1/2}))^{\mathrm{cl}} = (2nL(B_{1/2}))^{\mathrm{cl}} = (2n)(L(B_{1/2}))^{\mathrm{cl}}$, and we see that $(L(B_{1/2}))^{\mathrm{cl}}$ contains some nonempty open subset V of Y. If v and v' are in V, they are in $(L(B_{1/2}))^{\mathrm{cl}}$ and there exist sequences $\{v_n\}$ and $\{v'_n\}$ in $L(B_{1/2})$ with $v_n \to v$ and $v'_n \to v'$. By linearity, $v_n - v'_n$ is in $L(B_1)$, and passage to the limit shows that $v - v'$ is in $L(B_1)^{\mathrm{cl}}$. The set $V - V$ of such differences $v - v'$ is the union over $v' \in V$ of $V - v'$, hence is the union of open sets and is open. Since 0 is in $V - V$, the set $V - V$ is an open neighborhood of 0 lying in $L(B_1)^{\mathrm{cl}}$.

The second step is to show that the image of any neighborhood of 0 in X is a neighborhood of 0 in Y. The previous step shows that $(L(B_1))^{\mathrm{cl}} \supseteq U_s$ for some $s > 0$, and we show for every $c > 0$ that $L(B_c) \supseteq U_{sc/2}$. For $t > 0$, multiplication of the inclusion $(L(B_1))^{\mathrm{cl}} \supseteq U_s$ by t shows that

$$(L(B_t))^{\mathrm{cl}} \supseteq U_{st} \tag{$*$}$$

since multiplication by t is a homeomorphism of Y and L is linear. If y is in $U_{sc/2}$, we are to produce x in B_c with $L(x) = y$, and we do so by successive approximations. Specifically we construct inductively the terms x_n of a convergent series in X with sum x, as follows: Condition $(*)$ with $t = c/2$ allows us to choose a member x_1 of $B_{c/2}$ with $\|y - L(x_1)\| < 2^{-2}sc$. If $x_1, \dots, x_{n-1}$ have been constructed with each x_j in $B_{2^{-j}c}$ and with

$$\|y - L(x_1 + \cdots + x_{n-1})\| < 2^{-n}sc,$$

then $y - L(x_1 + \cdots + x_{n-1})$ is in $U_{2^{-n}sc}$. Condition $(*)$ with $t = 2^{-n}c$ shows that we can find x_n in $B_{2^{-n}c}$ with

$$\|y - L(x_1 + \cdots + x_{n-1}) - L(x_n)\| < 2^{-(n+1)}sc.$$

We now have

$$\|y - L(x_1 + \cdots + x_{n-1} + x_n)\| < 2^{-(n+1)}sc.$$

This completes the inductive construction of the x_n's, and we shall prove that the series $\sum x_n$ is convergent in X. Since X is complete, it is enough to show that the partial sums of $\sum x_n$ are Cauchy. If $q \geq p$, then

$$\big\| \textstyle\sum_{n=1}^q x_n - \sum_{n=1}^p x_n \big\| = \big\| \sum_{n=p+1}^q x_n \big\| \leq \sum_{n=p+1}^q \|x_n\| \leq \sum_{n=p+1}^q 2^{-n}c.$$

The right side is $\leq 2^{-p}c$, and the partial sums of $\sum x_n$ are indeed Cauchy. Let $x = \sum_{n=1}^\infty x_n$. Taking $p = 0$ and using the continuity of the norm, we see that $\|x\| \leq c$. By continuity of L, we have $y = \lim_n L(x_1 + \cdots + x_n) = L(x)$. Consequently the member y of $U_{sc/2}$ is of the form $L(x)$ for some x in B_c, as was asserted.

The third step is to show that each open set of X is mapped to an open set of Y by L. Let U be open in X, let x be in U, and let N be an open neighborhood of 0 in X such that $x + N \subseteq U$. The previous step shows that there is some open neighborhood V of 0 in Y such that $V \subseteq L(N)$. Then $L(x) + V$ is an open neighborhood in Y of $L(x)$ with

$$L(x) + V \subseteq L(x) + L(N) = L(x + N) \subseteq L(U).$$

Therefore $L(U)$ contains a neighborhood about each of its points and must be open.

Corollary 12.25. A one-one continuous linear operator L of a Banach space X onto a Banach space Y has a continuous linear inverse.

PROOF. Since L is one-one onto, L^{-1} exists. For L^{-1} to be continuous, the inverse image under L^{-1} of each open set is to be open. In other words, the direct image under L of any open set is to be open. But this is just the conclusion of Theorem 12.24.

EXAMPLE. Let $\mathcal{F}$ be the Fourier coefficient mapping, which carries functions in $L^1\big(\frac{1}{2\pi}\,dx\big)$ to doubly infinite sequences $\{c_n\}$ vanishing at infinity. The linear operator $\mathcal{F}$ has norm 1 when the space of doubly infinite sequences is given the supremum norm $\|\{c_n\}\|_{\text{sup}} = \sup_n |c_n|$. Corollary 6.50 shows that $\mathcal{F}$ is one-one. Let us see that there is some doubly infinite sequence vanishing at infinity that is not the sequence of Fourier coefficients of some L^1 function. If this were not so, then Corollary 12.25 would say that $\mathcal{F}^{-1}$ is bounded. We can obtain a contradiction if we produce a sequence $\{f_n\}$ of L^1 functions with $\|f_n\|_1 = 1$ for all n and with $\lim_n \|\mathcal{F}(f_n)\|_{\text{sup}} = 0$. Form the Dirichlet kernel D_n as defined in Section I.10 and reproduced in the previous section. Its Fourier coefficients c_k are 1 for $|k| \leq n$ and are 0 for $|k| > n$, and thus $\|\mathcal{F}(D_n)\|_{\text{sup}} = 1$. Put $f_n = D_n\big/\|D_n\|_1$. Then $\|f_n\|_1 = 1$ for all n, and $\|\mathcal{F}(f_n)\|_{\text{sup}} = 1\big/\|D_n\|_1$. Proposition 12.23 shows that in fact $\lim_n 1/\|D_n\|_1 = 0$, and we obtain the desired contradiction. The conclusion is that the image of $\mathcal{F}$ on L^1 fails to include some doubly infinite sequence $\{c_n\}$ vanishing at infinity.

If $f : X \to Y$ is a function between Hausdorff spaces, the graph of f is the subset $G = \{(x, f(x)) \mid x \in X\}$ of $X \times Y$. If f is continuous, then G is a closed set, as we see immediately by using nets. The converse fails because $f : [0, 1] \to \mathbb{R}$ with $f(0) = 0$ and $f(x) = 1/x$ for $x > 0$ is a discontinuous function with closed graph.

We shall be interested in the converse under the additional condition that our function f is linear. Our spaces being metric spaces, the condition that the graph be closed is that whenever $\{(x_n, f(x_n))\}$ converges to some (x, y), then x is in the domain of f and $f(x) = y$.

Linearity by itself is not enough to get an affirmative result. In fact, let $X = C([0, 1])$, let X_0 be the vector subspace of functions with a continuous derivative, and let $L : X_0 \to X$ be the derivative operator $F \mapsto F'$. If $\lim_n F_n = F$ in X and $\lim_n F_n' = H$, then Theorem 1.23 shows that F' exists and equals H. Hence the linear operator $L : X_0 \to X$ has closed graph. However, L is unbounded since the function $x \mapsto x^n$ has norm 1 and its derivative has norm n.

Corollary 12.26 (Closed Graph Theorem). If $L : X \to Y$ is a linear operator from a Banach space X into a Banach space Y such that the graph of L is a closed subset of $X \times Y$, then L is a bounded linear operator.

PROOF. Make $X \oplus Y$ into a Banach space by defining $\|(x, y)\| = \|x\|_X + \|y\|_Y$. The graph $G = \{(x, L(x)) \mid x \in X\}$ of L is a vector subspace of $X \oplus Y$ since L is linear, and it is closed by hypothesis. Thus G is a Banach space. The linear operator $P : G \to X$ given by $P((x, L(x)) = x$ is one-one and onto, and Corollary 12.25 shows that the linear operator $P^{-1} : X \to G$ given by $P^{-1}(x) = (x, L(x))$ is continuous. If E denotes the projection of $X \oplus Y$ to the Y coordinate, then E is bounded with norm ≤ 1, and hence the restriction $E\big|_G : G \to Y$ is bounded with norm ≤ 1. Therefore the composition $(E\big|_G) \circ P^{-1} : X \to Y$ is bounded. But $(E\big|_G)(P^{-1}(x)) = E(x, L(x)) = L(x)$, and thus L is bounded.

EXAMPLE. Suppose that a Banach space X is the vector-space direct sum of two closed vector subspaces: $X = Y \oplus Y'$. Let $E : X \to Y$ be the projection of X on Y given by $E(y + y') = y$. Corollary 12.26 implies that E is bounded. In fact, let $x_n = y_n + y_n'$ define a sequence in X, so that (x_n, y_n) defines a sequence in the graph of E. Suppose that $\lim_n(x_n, y_n) = (x_0, y_0)$ in $X \times X$, i.e., that $\lim_n x_n = x_0$ and $\lim_n y_n = y_0$. Here x_0 is in X, and y_0 is in Y since Y is closed. Then $y_0' = \lim_n y_n' = \lim_n x_n - \lim_n y_n = x_0 - y_0$, and this is in Y' since Y' is closed. The equality $x_0 = y_0 + y_0'$ shows that $E(x_0) = y_0$, and therefore (x_0, y_0) is in the graph of E. In other words, the graph of E is closed. We conclude from Corollary 12.26 that E is bounded.

7. Problems

1. Let X be a normed linear space.
   (a) Prove that the closure of the open ball of radius r and center x_0 is the closed ball of radius r and center x_0.
   (b) If X is complete, prove that any decreasing sequence of closed balls has nonempty intersection.
2. The normed linear space $C^{(N)}([a, b])$ was defined in Section 1. Prove that it is complete.
3. The normed linear space $H^\infty(D)$ and its vector subspace $A(D)$ were defined in Section 1. Prove that $H^\infty(D)$ is complete and that $A(D)$ is a closed subspace, hence complete.
4. Let X be a Banach space, let Y be a closed vector subspace, and define $\|x+Y\| = \inf_{y\in Y} \|x + y\|$ for $x + Y$ in the quotient vector space X/Y.
   (a) Show that $\| \cdot +Y\|$ is a norm for X/Y.
   (b) By replacing a Cauchy sequence $\{x_n + Y\}$ in X/Y by a subsequence such that $\|x_{n_k} - x_{n_{k+1}} + Y\| \le 2^{-k}$, show that the subsequence can be lifted to a Cauchy sequence in X and deduce that X/Y is a Banach space.
5. Let $v_1, \dots, v_n$ be vectors in an inner-product space. Their **Gram matrix** is the Hermitian matrix of inner products given by $G(v_1, \dots, v_n) = [(v_i, v_j)]$, and $\det G(v_1, \dots, v_n)$ is called their **Gram determinant**.
   (a) If $c_1, \dots, c_n$ are in $\mathbb{C}$, let $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. Prove that $c^{\mathrm{tr}} G(v_1, \dots, v_n)\bar{c} = \|c_1 v_1 + \cdots + c_n v_n\|^2$.
   (b) Making use of the finite-dimensional Spectral Theorem, prove that there exists a unitary matrix u such that the matrix $u^{-1}G(v_1, \dots, v_n)u$ is diagonal with diagonal entries ≥ 0.
   (c) Prove that $\det G(v_1, \dots, v_n) \ge 0$ with equality if and only if $v_1, \dots, v_n$ are linearly dependent. (This generalizes the Schwarz inequality.)
6. **(Gram–Schmidt orthogonalization process)** Let $(u_1, \dots, u_n)$ be a linearly independent ordered set in an inner-product space, and inductively define $v'_1 = u_1$, $v_1 = \|v'_1\|^{-1}v'_1$, $v'_k = u_k - \sum_{j=1}^{k-1} (u, v_j)v_j$, and $v_k = \|v'_k\|^{-1}v'_k$. Prove that the vectors $v_1, \dots, v_n$ are well defined, that $v_1, \dots, v_n$ are orthonormal, and that for each k with $1 \le k \le n$, $\operatorname{span}\{v_1, \dots, v_k\} = \operatorname{span}\{u_1, \dots, u_k\}$.
7. Let H_1 and H_2 be Hilbert spaces with respective orthonormal bases $\{u_\alpha\}$ and $\{v_\beta\}$. If there is a one-one function carrying the one orthonormal basis onto the other, prove that there is a bounded linear operator $F : H_1 \to H_2$ carrying H_1 onto H_2 and preserving distances. Deduce that H_1 and H_2 are isomorphic as Hilbert spaces if and only if they have the same Hilbert space dimension.

8. Let (S, μ) be a σ-finite measure space, and let f be in $L^\infty(S, \mu)$.
   (a) Show that multiplication by f is a bounded linear operator on $L^2(S, \mu)$, and find the norm of this operator.
   (b) Find the adjoint of the operator in (a).

9. Suppose that X is a normed linear space and that its dual X^* is separable in its norm topology, with $\{x_n^*\}$ as a countable dense set. For each n, choose x_n in X with $\|x_n\| \leq 1$ and $|x_n^*(x_n)| \geq \frac{1}{2}\|x_n^*\|$. Prove that $\{x_n\}$ is dense in X, so that X^* separable implies X separable.

10. By considering the discontinuous indicator function $I_{\{s_0\}}$, where s_0 is a limit point of S, prove that the Banach space $C(S)$ is not reflexive if S is compact Hausdorff and infinite.

11. Without using the Baire Category Theorem, prove that the Uniform Boundedness Theorem for linear functionals implies the same theorem for linear operators.

12. Suppose for each n that $L_n : X \to X'$ is a bounded linear operator from a normed linear space X to a Banach space X' such that $\|L_n\| \leq C$ with C independent of n. Suppose in addition that $\{L_n(y)\}$ converges for each y in a dense subset Y of X. Prove that $L(x) = \lim_n L_n(x)$ exists for all x in X and that the resulting function $L : X \to X'$ is a bounded linear operator with $\|L\| \leq C$.

13. Let X be a normed linear space, and let $\{x_\alpha\}$ be a subset of X. If $\sup_\alpha |x^*(x_\alpha)| < \infty$ for each x^* in X^*, prove that $\sup_\alpha \|x_\alpha\| < \infty$.

14. Let X be a Banach space. A subset E of X is **convex** if it contains all points $(1-t)x + ty$ with $0 \leq t \leq 1$ whenever it contains x and y.
    (a) Show that any closed ball $\{y \mid |y - x| \leq r\}$ is convex.
    (b) Give an example of a decreasing sequence of nonempty bounded closed convex sets in a Banach space with empty intersection.

15. Let X and Y be Banach spaces, and let L be a bounded linear operator from X onto Y. Suppose that $\{y_n\}$ is a convergent sequence in Y with limit y_0. Prove that there exists a constant M and a sequence $\{x_n\}$ in X such that $\|x_n\| \leq M\|y_n\|$ for all n, $L(x_n) = y_n$ for all n, and $\{x_n\}$ is convergent.

Problems 16–18 introduce "Banach limits," a kind of universal summability method. Let X be the real Banach space of real-valued bounded sequences $s = \{s_n\}_{n=1}^\infty$ with the supremum norm.

16. Let X_0 be the smallest closed vector subspace of X containing all sequences with terms $s_1, s_2 - s_2, s_3 - s_2, \ldots$ such that $\{s_n\}$ is in X. Prove that the sequence e with all terms 1 is not in X_0.

17. A **Banach limit** is defined to be any member x^* of X^* with $\|x^*\| = 1$, $x^*(e) = 1$, and $x^*(x_0) = 0$ for all x_0 in X_0. Prove that a Banach limit exists.

18. Let $\mathrm{LIM}_{n\to\infty} s_n$ denote the value of a Banach limit when applied to the member $\{s_n\}$ of X. Prove that this satisfies
   (a) $\mathrm{LIM}_{n\to\infty} s_n \ge 0$ if $s_n \ge 0$ for all n.
   (b) $\mathrm{LIM}_{n\to\infty} s_{n+1} = \mathrm{LIM}_{n\to\infty} s_n$ for every $\{s_n\}$ in X.
   (c) $\mathrm{LIM}_{n\to\infty} s_n = 0$ if all terms s_n are 0 for n sufficiently large.
   (d) $\liminf_n s_n \le \mathrm{LIM}_{n\to\infty} s_n \le \limsup_n s_n$ for all $\{s_n\}$ in X.
   (e) $\mathrm{LIM}_{n\to\infty} s_n = c$ if $\{s_n\}$ is convergent with limit c.

Problems 19–24 establish the Jordan and von Neumann Theorem that a normed linear space satisfying the parallelogram law acquires its norm from an inner product, the definition of the inner product being $(x, y) = \sum_k \frac{i^k}{4}\|x + i^k y\|^2$, where the sum extends for $k \in \{0, 2\}$ if the scalars are real and extends for $k \in \{0, 1, 2, 3\}$ if the scalars are complex. The norm is recovered from the inner product by the usual formula $(x, x) = \|x\|^2$. Thus let X be a normed linear space with norm $\|\cdot\|$ such that the parallelogram law holds.

19. Check from the definition of (x, y) that $(x, x) = \|x\|^2$, that $(x, x) \ge 0$ with equality if and only if $x = 0$, and that $(x, y) = \overline{(y, x)}$.

20. Prove the identity
$$\|x + y + z\|^2 = \|x + y\|^2 + \|x + z\|^2 + \|y + z\|^2 - \|x\|^2 - \|y\|^2 - \|z\|^2$$
for all x, y, z in X.

21. Derive the formula $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$ from the identity in the previous problem.

22. Let D be the set of rationals if the scalars are real, or the set of all $a + bi$ with a and b rational if the scalars are complex. Using the definition of (x, y) and the result of the previous problem, prove that $(rx, y) = r(x, y)$ if r is in D.

23. By considering $\|x - ry\|^2$ for r in D with r tending to $(x, y)/\|y\|^2$, prove that $(\cdot\,, \cdot)$ satisfies the Schwarz inequality.

24. By estimating $|r(x, y) - (cx, y)|$ with the Schwarz inequality when c is a scalar and r is a member of D tending to c, prove that $c(x, y) = (cx, y)$, thereby completing the proof that $(\cdot\,, \cdot)$ is an inner product.

APPENDIX

Abstract. This appendix treats some topics that are likely to be well known by some readers and less well known by others. Section A1 deals with set theory and with functions: it discusses the role of formal set theory, it works in a simplified framework that avoids too much formalism and the standard pitfalls, it establishes notation, and it mentions some formulas. Some emphasis is put on distinguishing the image and the range of a function, as this distinction is important in algebra and algebraic topology and therefore plays a role when real analysis begins to interact seriously with algebra.

Sections A2 and A3 assume knowledge of Section I.1 and discuss topics that occur logically between the end of Section I.1 and the beginning of Section I.2. The first of these establishes the Mean Value Theorem and its standard corollaries and then goes on to define the notion of a continuous derivative for a function on a closed interval. The other section gives a careful treatment of the differentiability of an inverse function in one-variable calculus.

Section A4 is a quick review of complex numbers, real and imaginary parts, complex conjugation, and absolute value. Complex-valued functions appear in the book beginning in Section I.5. Section A5 states and proves the classical Schwarz inequality, which is used in Chapter II to establish the triangle inequality for certain metrics but is needed before that in Chapter I in the context of Fourier series.

Sections A6 and A7 are not needed until Chapter II. The first of these defines equivalence relations and establishes the basic fact that they lead to a partitioning of the underlying set into equivalence classes. The other section discusses the connection between linear functions and matrices in the subject of linear algebra and summarizes the basic properties of determinants.

Section A8, which is not needed until Chapter IV, establishes unique factorization for polynomials with real or complex coefficients and defines "multiplicity" for roots of complex polynomials.

Sections A9 and A10 return to set theory. Section A9 defines partial orderings and includes Zorn's Lemma, which is a powerful version of the Axiom of Choice, while Section A10 concerns cardinality. The material in these sections first appears in problems in Chapter V; it does not appear in the text until Chapter X in the case of Section A9 and until Chapter XII in the case of Section A10.

A1. Sets and Functions

Real analysis typically makes use of an informal notion of set theory and notation for it in which sets are described by properties of their elements and by operations on sets. This informal set theory, if allowed to be too informal, runs into certain paradoxes, such as the **Russell paradox:** "If S is the set of all sets that do not contain themselves as elements, is S a member of S or is it not?" The conclusion of the Russell paradox is that the "set" of all sets that do not contain themselves as elements is not in fact a set.

Mathematicians' experience is that such pitfalls can be avoided completely by working within some formal axiom system for sets, of which there are several that are well established. A basic one is "Zermelo–Fraenkel set theory," and the remarks in this section refer specifically to it but refer to the others at least to some extent.[1]

The standard logical paradoxes are avoided by having sets, elements (or "entities"), and a membership relation $\in$ such that $a \in S$ is a meaningful statement, true or false, if and only if a is an element and S is a set. The terms **set**, **element**, and $\in$ are taken to be primitive terms of the theory that are in effect defined by a system of axioms. The axioms ensure the existence of many sets, including infinite sets, and operations on sets that lead to other sets. To make full use of this axiom system, one has to regard it as occurring in the context of certain rules of logic that tell the forms of basic statements (namely, $a = b$, $a \in S$, and "S is a set"), the connectives for creating complicated statements from simple ones ("or," "and," "not," and "if ... then"), and the way that quantifiers work ("there exists" and "for all").

Working rigorously with such a system would likely make the development of mathematics unwieldy, and it might well obscure important patterns and directions. In practice, therefore, one compromises between using a formal axiom system and working totally informally; let us say that one works "informally but carefully." The logical problems are avoided not by rigid use of an axiom system, but by taking care that sets do not become too "large": one limits the sets that one uses to those obtained from other sets by set-theoretic operations and by passage to subsets.[2]

A feature of the axiom system that one takes advantage of in working informally but carefully is that the axiom system does not preclude the existence of additional sets beyond those forced to exist by the axioms. Thus, for example, in the subject of coin-tossing within probability, it is normal to work with the set of possible outcomes as $S = \{\text{heads, tails}\}$ even though it is not apparent that requiring this S to be a set does not introduce some contradiction.

It is worth emphasizing that the points of the theory at which one takes particular care vary somewhat from subject to subject within mathematics. For example, it is sometimes of interest in calculus of several variables to distinguish between the range of a function and its image in a way that will be mentioned below, but it is usually not too important. In homological algebra, however, the distinction is

[1]Mathematicians have no proof that this technique avoids problems completely. Such a proof would be a proof of the consistency of a version of mathematics in which one can construct the integers, and it is known that this much of mathematics cannot be proved to be consistent unless it is in fact inconsistent.

[2]Not every set so obtained is to be regarded as "constructed." The Axiom of Choice, which we come to shortly, is an existence statement for elements in products of sets, and the result of applying the axiom is a set that can hardly be viewed as "constructed."

extremely important, and the subject loses a great deal of its impact if one blurs the notions of range and image.

Some references for set theory that are appropriate for reading *once* are Halmos's *Naive Set Theory*, Hayden–Kennison's *Zermelo–Fraenkel Set Theory*, and Chapter 0 and the appendix of Kelley's *General Topology*. The Kelley book is one that uses the word "class" as a primitive term more general than "set"; it develops von Neumann set theory.

All that being said, let us now introduce the familiar terms, constructions, and notation that one associates with set theory. To cut down on repetition, one allows some alternative words for "set," such as **family** and **collection**. The word "class" is used by some authors as a synonym for "set," but the word **class** is used in some set-theory axiom systems to refer to a more general notion than "set," and it will be useful to preserve this possibility. Thus a class can be a set, but we allow ourselves to speak, for example, of the class of all groups even though this class is too large to be a set. Alternative terms for "element" are **member** and **point**; we shall not use the term "entity." Instead of writing $\in$ systematically, we allow ourselves to write "in." Generally, we do not use $\in$ in sentences of text as an abbreviation for an expression like "is in" that contains a verb.

If A and B are two sets, some familiar operations on them are the **union** $A \cup B$, the **intersection** $A \cap B$, and the **difference** $A - B$, all defined in the usual way in terms of the elements they contain. Notation for the difference of sets varies from author to author; some other authors write $A \setminus B$ or $A \sim B$ for difference, but this book uses $A - B$. If one is thinking of A as a universe, one may abbreviate $A - B$ as B^c, the **complement** of B in A. The empty set $\varnothing$ is a set, and so is the set of all subsets of a set A, which is sometimes denoted by 2^A. Inclusion of a subset A in a set B is written $A \subseteq B$ or $B \supseteq A$. Inclusion that does not permit equality is denoted by $A \subsetneqq B$ or $B \supsetneqq A$; in this case one says that A is a **proper subset** of B or that A is **properly contained** in B.

If A is a set, the **singleton** $\{A\}$ is a set with just the one member A. Another operation is **unordered pair**, whose formal definition is $\{A, B\} = \{A\} \cup \{B\}$ and whose informal meaning is a set of two elements in which we cannot distinguish either element over the other. Still another operation is **ordered pair**, whose formal definition is $(A, B) = \{\{A\}, \{A, B\}\}$. It is customary to think of an ordered pair as a set with two elements in which one of the elements can be distinguished as coming first.[3]

Let A and B be two sets. The set of all ordered pairs of an element of A and

[3]Unfortunately a "sequence" as in Chapter I gets denoted by $\{x_1, x_2, \dots\}$ or $\{x_n\}_{n=1}^{\infty}$. If its notation were really consistent with the above definitions, we might infer, inaccurately, that the order of the terms of the sequence does not matter. The notation for unordered pairs, ordered pairs, and sequences is, however, traditional, and it will not be changed here.

an element of B is a set denoted by $A \times B$; it is called the **product** of A and B or the **Cartesian product**. A **relation** between a set A and a set B is a subset of $A \times B$. Functions, which are to be defined in a moment, provide examples. Two examples of relations that are usually not functions are "equivalence relations," which are discussed in Section A6, and "partial orderings," which are discussed in Section A9.

If A and B are sets, a relation f between A and B is said to be a **function**, written $f : A \to B$, if for each $x \in A$, whenever $y \in B$ and $z \in B$ are such that (x, y) and (x, z) are in f, then $y = z$. If (x, y) is in f, we write $f(x) = y$. In this informal but careful definition of function, the function consists of more than just a set of ordered pairs; it consists of the set of ordered pairs regarded as a subset of $A \times B$. This careful definition makes it meaningful to say that the set A is the **domain**, the set B is the **range**, and the subset of $y \in B$ such that $y = f(x)$ for some $x \in A$ is the **image** of f. The image is also denoted by $f(A)$. Sometimes a function f is described in terms of what happens to typical elements, and then the notation is $x \mapsto f(x)$ or $x \mapsto y$, possibly with y given by some formula or by some description in words about how it is obtained from x. Sometimes a function f is written as $f(\,\cdot\,)$, with a dot indicating the placement of the variable; this notation is especially helpful in working with restrictions of functions, which we come to in a moment, and with functions of two variables when one of the variables is held fixed. This notation is useful also for functions that involve unusual symbols, such as the absolute value function $x \mapsto |x|$, which in this notation becomes $|\cdot|$. The word **map** or **mapping** is sometimes used for "function" and for the operation of a function, particularly when a geometric context for the function is of importance.

Often mathematicians are not so careful with the definition of function. Depending on the degree of informality that is allowed, one may occasionally refer to a function as $f(x)$ when it should be called f or $x \mapsto f(x)$. If any confusion is possible, it is wise to use the more rigorous notation. Another habit of informality is to regard a function $f : A \to B$ as simply a set of ordered pairs. Thus two functions $f_1 : A \to B$ and $f_2 : A \to C$ become the same if $f_1(a) = f_2(a)$ for all a in A. With the less careful definition, the notion of the range of a function is not really well defined. The less careful definition can lead to trouble in algebra, but it does not often lead to trouble in real analysis until one gets to a level where algebra and analysis merge somewhat.

The set of all functions from a set A to a set B is a set. It is sometimes denoted by B^A. The special case 2^A that arose with subsets comes by regarding 2 as a set $\{1, 2\}$ and identifying a function f from A into $\{1, 2\}$ with the subset of all elements x of A for which $f(x) = 1$.

If a subset B of a set A may be described by some distinguishing property P of its elements, we may write this relationship as $B = \{x \in A \mid P\}$. For example, the function f in the previous paragraph is identified with the subset $\{x \in A \mid f(x) = 1\}$. Another example is the image of a general function $f : A \to B$, namely $f(A) = \{y \in B \mid y = f(x) \text{ for some } x \in A\}$. Still more generally along these lines, if E is any subset of A, then $f(E)$ denotes the set $\{y \in B \mid y = f(x) \text{ for some } x \in E\}$. Some authors use a colon instead of a vertical line in this notation.

This book frequently uses sets denoted by expressions like $\bigcup_{x \in S} A_x$, an indexed union, where S is a set that is usually nonempty. If S is the set $\{1, 2\}$, this reduces to $A_1 \cup A_2$. In the general case it is understood that we have an unnamed function, say f, given by $x \mapsto A_x$, having domain S and range an unnamed set T, and $\bigcup_{x \in S} A_x$ is the set of all $y \in T$ such that y is in A_x for some $x \in S$. When S is understood, we may write $\bigcup_x A_x$ instead of $\bigcup_{x \in S} A_x$. Indexed intersections $\bigcap_{x \in S} A_x$ are defined similarly, and this time it is essential to disallow S empty because otherwise the intersection cannot be a set in any useful set theory.

There is also an indexed **Cartesian product** $\bigtimes_{x \in S} A_x$ that specializes in the case that $S = \{1, 2\}$ to $A_1 \times A_2$. Usually S is assumed nonempty. This Cartesian product is the set of all functions f from S into $\bigcup_{x \in S} A_x$ such that $f(x)$ is in A_x for all $x \in S$. In the special case that S is $\{1, \dots, n\}$, the Cartesian product is the set of ordered n-tuples from n sets $A_1, \dots, A_n$ and may be denoted by $A_1 \times \cdots \times A_n$; its members may be denoted by $(a_1, \dots, a_n)$ with $a_j \in A_j$ for $1 \le j \le n$. When the factors of a Cartesian product have some additional algebraic structure, the notation for the Cartesian product is sometimes altered; for example, the Cartesian product of groups A_x is denoted by $\prod_{x \in S} A_x$.

It is completely normal in real analysis, and it is the practice in this book, to take the following axiom as part of one's set theory; the axiom is normally used without specific mention.

Axiom of Choice. The Cartesian product of nonempty sets is nonempty.

If the index set is finite, then the Axiom of Choice reduces to a theorem of set theory. The axiom is often used quite innocently with a countably infinite index set. For example, Proposition 1.7c asserts that any sequence in $\mathbb{R}^*$ has a subsequence converging to $\limsup a_n$, and the proof constructs one member of the sequence at a time. When these members have some flexibility in their definitions, as is the case with the proof as it is written for Proposition 1.7c, the Axiom of Choice is being invoked. When the members instead have specific definitions, such as "the term a_n such that n is the smallest integer satisfying such-and-such properties," the axiom is not being invoked. The proof in the text of Proposition 1.7c can be rewritten with specific definitions and thereby can avoid invoking the axiom, but there is no point in undertaking this rewriting. In Chapter II the axiom is invoked in situations in which the index set is uncountable; uses of compactness provide a number of examples.

From the Axiom of Choice, one can deduce a powerful tool known as Zorn's Lemma, whose use it is normal to acknowledge. Zorn's Lemma appears in Section A9 and is used in problems beginning in Chapter V and in the text beginning in Chapter X.

If $f : A \to B$ is a function and B is a subset of B', then f can be regarded as a function with range B' in a natural way. Namely, the set of ordered pairs is unchanged but is to be regarded as a subset of $A \times B'$ rather than $A \times B$.

Let $f : A \to B$ and $g : B \to C$ be two functions such that the range of f equals the domain of g. The **composition** $g \circ f : A \to C$ is the function with $(g \circ f)(x) = g(f(x))$ for all x. Because of the construction in the previous paragraph, it is meaningful to define the composition more generally when the range of f is merely a subset of the domain of g.

A function $f : A \to B$ is said to be **one-one** if $f(x_1) \neq f(x_2)$ whenever x_1 and x_2 are distinct members of A. The function is said to be **onto**, or often "onto B," if its image equals its range. The terminology "onto B" avoids confusion: it specifies the image and thereby guards against the use of the less careful definition of function mentioned above. A mathematical audience often contains some people who use the careful definition of function and some people who use the less careful definition. For the latter kind of person, a function is always onto something, namely its image, and a statement that a particular function is onto might be regarded as a tautology.

When a function $f : A \to B$ is one-one and is onto B, there exists a function $g : B \to A$ such that $g \circ f$ is the identity function on A and $f \circ g$ is the identity function on B. The function g is unique, and it is defined by the condition, for $y \in B$, that $g(y)$ is the unique $x \in A$ with $f(x) = y$. The function g is called the **inverse function** of f and is often denoted by f^{-1}.

Conversely if $f : A \to B$ has an inverse function, then f is one-one and is onto B. The reason is that a composition $g \circ f$ can be one-one only if f is one-one, and in addition, that a composition $f \circ g$ can be onto the range of f only if f is onto its range.

If $f : A \to B$ is a function and E is a subset of A, the **restriction** of f to E, denoted by $f\big|_E$, is the function $f : E \to B$ consisting of all ordered pairs $(x, f(x))$ with $x \in E$, this set being regarded as a subset of $E \times B$, not of $A \times B$. One especially common example of a restriction is restriction to one of the variables of a function of two variables, and then the idea of using a dot in place of a variable can be helpful notationally. Thus the function of two variables might be indicated by f or $(x, y) \mapsto f(x, y)$, and the restriction to the first variable, for fixed value of the second variable, would be $f(\,\cdot\,, y)$ or $x \mapsto f(x, y)$.

We conclude this section with a discussion of direct and inverse images of sets under functions. If $f : A \to B$ is a function and E is a subset of A, we have defined $f(E) = \{y \in B \mid y = f(x) \text{ for some } x \in E\}$. This is the same as the image of $f\big|_E$ and is frequently called the image or **direct image** of E under f. The notion of direct image does not behave well with respect to some set-theoretic operations: it respects unions but not intersections. In the case of unions, we have

$$f\Big(\bigcup_{s\in S} E_s\Big) = \bigcup_{s\in S} f(E_s);$$

the inclusion $\supseteq$ follows since $f\Big(\bigcup_{s\in S} E_s\Big) \supseteq f(E_s)$ for each s, and the inclusion $\subseteq$ follows because any member of the left side is f of a member of some E_s. In the case of intersections, the question $f(E\cap F) \overset{?}{=} f(E)\cap f(F)$ can easily have a negative answer, the correct general statement being $f(E\cap F) \subseteq f(E)\cap f(F)$. An example with equality failing occurs when $A = \{1, 2, 3\}$, $B = \{1, 2\}$, $f(1) = f(3) = 1$, $f(2) = 2$, $E = \{1, 2\}$ and $F = \{2, 3\}$ because $f(E\cap F) = \{2\}$ and $f(E)\cap f(F) = \{1, 2\}$.

If $f : A \to B$ is a function and E is a subset of B, the **inverse image** of E under f is the set $f^{-1}(E) = \{x \in A \mid f(x) \in E\}$. This is well defined even if f does not have an inverse function. (If f does have an inverse function f^{-1}, then the inverse image of E under f coincides with the direct image of E under f^{-1}.)

Unlike direct images, inverse images behave well under set-theoretic operations. If $f : A \to B$ is a function and $\{E_s \mid s \in S\}$ is a set of subsets of B, then

$$f^{-1}\Big(\bigcap_{s\in S} E_s\Big) = \bigcap_{s\in S} f^{-1}(E_s),$$
$$f^{-1}\Big(\bigcup_{s\in S} E_s\Big) = \bigcup_{s\in S} f^{-1}(E_s),$$
$$f^{-1}(E_s^c) = (f^{-1}(E_s))^c.$$

In the third of these identities, the complement on the left side is taken within B, and the complement on the right side is taken within A. To prove the first identity, we observe that $f^{-1}\big(\bigcap_{s\in S} E_s\big) \subseteq f^{-1}(E_s)$ for each $s \in S$ and hence $f^{-1}\big(\bigcap_{s\in S} E_s\big) \subseteq \bigcap_{s\in S} f^{-1}(E_s)$. For the reverse inclusion, if x is in $\bigcap_{s\in S} f^{-1}(E_s)$, then x is in $f^{-1}(E_s)$ for each s and thus $f(x)$ is in E_s for each s. Hence $f(x)$ is in $\bigcap_{s\in S} E_s$, and x is in $f^{-1}\big(\bigcap_{s\in S} E_s\big)$. This proves the reverse inclusion. The second and third identities are proved similarly.

A2. Mean Value Theorem and Some Consequences

This section states and proves the Mean Value Theorem and two standard corollaries, and then it discusses the notion of a function with a continuous derivative on a closed interval. It makes use of results in Section I.1 of the text.

Lemma. Let $[a, b]$ be a nontrivial closed interval, and let $f : [a, b] \to \mathbb{R}$ be a continuous function that is differentiable on (a, b) and has $f(a) = f(b) = 0$. Then the derivative f' satisfies $f'(c) = 0$ for some c with $a < c < b$.

PROOF. We divide matters into three cases. If $f(x) > 0$ for some x, let c be a member of $[a, b]$ where f attains its maximum (existence by Theorem 1.11). Since $f(x) > 0$ somewhere, we must have $a < c < b$. Thus $f'(c)$ exists. If $f'(c) > 0$, then the inequality $\lim_{h\to 0} h^{-1}(f(c+h) - f(c)) > 0$ forces $f(c+h) > f(c)$ for h positive and sufficiently small, in contradiction to the fact that f attains its maximum at c. Similarly if $f'(c) < 0$, then we find that $f(c-h) > f(c)$ for h positive and sufficiently small, and again we have a contradiction. We conclude that $f'(c) = 0$.

If $f(x) \le 0$ for all x and $f(x) < 0$ for some x, let c instead be a member of $[a, b]$ where f attains its minimum. Arguing in the same way as in the previous paragraph, we find that $f'(c) = 0$.

Finally if $f(x) = 0$ for all x, then $f'(x) = 0$ for $a < x < b$, and $f'(c) = 0$ for $c = \frac{1}{2}(a + b)$, for example.

Mean Value Theorem. Let $[a, b]$ be a nontrivial closed interval. If $f : [a, b] \to \mathbb{R}$ is a continuous function that is differentiable on (a, b), then

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

for some c with $a < c < b$.

PROOF. Apply the lemma to the function

$$g(x) = f(x) - f(a) - (x - a)\,\tfrac{f(b)-f(a)}{b-a},$$

which has $g(a) = g(b) = 0$ and $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$.

Corollary 1. A differentiable function $f : (a, b) \to \mathbb{R}$ whose derivative is 0 everywhere on (a, b) is a constant function.

PROOF. If $f(a') \ne f(b')$, then the Mean Value Theorem produces some c between a' and b' where $f'(c) \ne 0$.

Corollary 2. A differentiable function $f : (a, b) \to \mathbb{R}$ whose derivative is > 0 everywhere on (a, b) is strictly increasing on (a, b).

PROOF. If $a' < b'$ and $f(a') \ge f(b')$, then the Mean Value Theorem produces some c with $a' < c < b'$ where $f'(c) \le 0$.

In the setting of the Mean Value Theorem, it can happen that $f'(x)$ has a **finite limit** C **as** x **decreases to** a (or as x increases to b). This terminology means that for any $\epsilon > 0$, there exists some $\delta > 0$ such that $|f'(x) - C| < \epsilon$ whenever $a < x < a+\delta$. In this case, f can be extended to a function F defined and continuous on $(-\infty, b]$, differentiable on $(-\infty, b)$, in such a way that F' is continuous at a. In fact, the extended definition is

$$F(x) = \begin{cases} f(x) & \text{for } a \le x \le b, \\ f(a) + C(x-a) & \text{for } -\infty < x \le a. \end{cases}$$

To see that $F'(a)$ exists for the extended function F, let $\epsilon > 0$ be given and choose $\delta > 0$ such that $a < x < a+\delta$ implies $|f'(x) - C| < \epsilon$. If $a < x < a+\delta$, then the Mean Value Theorem gives

$$\frac{F(x) - F(a)}{x-a} = F'(c)$$

with $a < c < x < a+\delta$, and hence $\left|\frac{F(x)-F(a)}{x-a} - C\right| < \epsilon$. If $a - \delta < x < a$, then

$$\left|\frac{F(x) - F(a)}{x-a} - C\right| = \left|\frac{(f(a) + C(x-a)) - f(a)}{x-a} - C\right| = 0.$$

Thus $F'(a)$ exists and equals C. The definitions make $\lim_{x\to a} F'(x) = F'(a)$, and hence F' is continuous at a.

As a consequence of this construction, it makes sense to say that a continuous function $f : [a, b] \to \mathbb{R}$ with a derivative on (a, b) has a continuous derivative at one or both endpoints. This phrasing means that f' has a finite limit at the endpoint in question, and it is equivalent to say that f extends to a larger set so as to be differentiable in an open interval about the endpoint and to have its derivative be continuous at the endpoint.

A3. Inverse Function Theorem in One Variable

This section addresses one of the "further topics" mentioned at the end of Section I.1 and assumes knowledge of Section I.1 and some additional facts about continuity and differentiability of functions of a real variable. The topic is that of differentiability of inverse functions, the nub of the matter being continuity of the inverse function. The topic is one that is sometimes skipped in calculus courses and slighted in courses in real variable theory. Yet it is necessary for the development of one of the two functions exp and log, of one of the two functions sin and arcsin, and of one of the two functions tan and arctan unless actual constructions of both members of a pair are given. In principle the matter arises also with differentiation of the function $x^{1/q}$ on $(0, \infty)$, but the proposition of this section can be readily avoided in that case by explicit calculations.

Proposition. Let (a, b) be an open interval in $\mathbb{R}$, possibly infinite, and let $f : (a, b) \to \mathbb{R}$ be a function with a continuous everywhere-positive derivative. Then f is strictly increasing and has an interval (c, d), possibly infinite, as its image. The inverse function $g : (c, d) \to (a, b)$ exists and has a continuous derivative given by $g'(y) = 1/f'(g(y))$.

PROOF. The function f is strictly increasing as a corollary of the Mean Value Theorem, and its image is an interval (c, d) because of the Intermediate Value Theorem (Theorem 1.12). Being one-one and onto, f has an inverse function g, according to Section A1. Fix $y_0 \in (c, d)$, fix $c'$and d' such that $c < c' < y_0 < d' < d$, and consider $y \neq y_0$ in (c', d'). Put $x = g(y)$, $x_0 = g(y_0)$, $a' = g(c')$, and $b' = g(d')$. Then $a < a' < x_0 < b' < b$ since f is strictly increasing.

By Theorem 1.11, there exist real numbers m and M such that $0 < m \leq f'(t) \leq M$ for all $t \in [a', b']$. The Mean Value Theorem produces ξ between x_0 and x such that

$$|y - y_0| = |f(x) - f(x_0)| = |f'(\xi)||x - x_0| \geq m|x - x_0|,$$

and hence $|x - x_0| \leq m^{-1}|y - y_0|$. Since g is one-one, we have $x \neq x_0$. Also, $f(x) = y \neq y_0 = f(x_0)$. Thus it makes sense to form

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)}.$$

Let $\epsilon > 0$ be given. Since $\lim_{t \to x_0} \frac{f(t) - f(x_0)}{t - x_0} = f'(x_0) \neq 0$, we have

$$\lim_{t \to x_0} \frac{t - x_0}{f(t) - f(x_0)} = \frac{1}{f'(x_0)}.$$

Choose $\eta > 0$ such that

$$\Big| \frac{t - x_0}{f(t) - f(x_0)} - \frac{1}{f'(x_0)} \Big| < \epsilon$$

as long as $|t - x_0| < \eta$ with $t \neq x_0$ and $t \in [a', b']$. Then put $\delta = \eta m$. If $|y - y_0| < \delta$, then $|x - x_0| \leq m^{-1}|y - y_0| < m^{-1}\delta = \eta$. Since $t = x$ satisfies the condition $|t - x_0| < \eta$ with $t \neq x_0$ and $t \in [a', b']$, it follows that

$$\Big| \frac{g(y) - g(y_0)}{y - y_0} - \frac{1}{f'(x_0)} \Big| = \Big| \frac{x - x_0}{f(x) - f(x_0)} - \frac{1}{f'(x_0)} \Big| < \epsilon$$

whenever $|y - y_0| < \delta$. Since ϵ is arbitrary, the conclusion is that $g'(y_0) = 1/f'(g(y_0))$. Since g is differentiable, g is continuous and also the composition $f' \circ g$ is continuous. Because $f' \circ g$ is nowhere zero, $g' = 1/(f' \circ g)$ is continuous. This completes the proof.

A4. Complex Numbers

Complex numbers are taken as known, and this section reviews their notation and basic properties.

Briefly, the system $\mathbb{C}$ of complex numbers is a two-dimensional vector space over $\mathbb{R}$ with a distinguished basis $\{1, i\}$ and a multiplication defined initially by $11 = 1$, $1i = i1 = i$, and $ii = -1$. Elements may then be written as $a + bi$ or $a + ib$ with a and b in $\mathbb{R}$; here a is an abbreviation for $a1$. The multiplication is extended to all of $\mathbb{C}$ so that the distributive laws hold, i.e., so that $(a+bi)(c+di)$ can be expanded in the expected way. The multiplication is associative and commutative, the element 1 acts as a multiplicative identity, and every nonzero element has a multiplicative inverse: $(a + bi)\big(\frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2}\big) = 1$.

Complex conjugation is indicated by a bar: the conjugate of $a + bi$ is $a - bi$ if a and b are real, and we write $\overline{a + bi} = a - bi$. Then we have $\overline{z + w} = \bar{z} + \bar{w}$, $\overline{rz} = r\bar{z}$ if r is real, and $\overline{zw} = \bar{z}\bar{w}$.

The **real** and **imaginary parts** of $z = a + bi$ are $\operatorname{Re} z = a$ and $\operatorname{Im} z = b$. These may be computed as $\operatorname{Re} z = \frac{1}{2}(z + \bar{z})$ and $\operatorname{Im} z = -\frac{i}{2}(z - \bar{z})$.

The **absolute value** function of $z = a + bi$ is given by $|z| = \sqrt{a^2 + b^2}$, and this satisfies $|z|^2 = z\bar{z}$. It has the simple properties that $|\bar{z}| = |z|$, $|\operatorname{Re} z| \le |z|$, and $|\operatorname{Im} z| \le |z|$. In addition, it satisfies

$$|zw| = |z||w|$$

because
$$|zw|^2 = zw\overline{zw} = zw\bar{z}\bar{w} = z\bar{z}w\bar{w} = |z|^2|w|^2,$$

and it satisfies the **triangle inequality**

$$|z + w| \le |z| + |w|$$

because
$$\begin{aligned} |z + w|^2 &= (z + w)\overline{(z + w)} = z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w} \\ &= |z|^2 + 2\operatorname{Re}(z\bar{w}) + |w|^2 \le |z|^2 + 2|z\bar{w}| + |w|^2 \\ &= |z|^2 + 2|z||w| + |w|^2 = (|z| + |w|)^2. \end{aligned}$$

A5. Classical Schwarz Inequality

The inequality in question is as follows.[4]

---

[4]In the classical setting below, the inequality is often called the "Cauchy–Schwarz inequality" and may have other people's names attached to it as well. However, generalizations tend to be called simply the "Schwarz inequality," and this book therefore drops all names but Schwarz.

Schwarz inequality. Let $(a_1, \dots, a_n)$ and $(b_1, \dots, b_n)$ be n-tuples of complex numbers. Then

$$\Big|\sum_{k=1}^{n} a_k \overline{b_k}\Big| \le \Big(\sum_{k=1}^{n} |a_k|^2\Big)^{1/2}\Big(\sum_{k=1}^{n} |b_k|^2\Big)^{1/2}.$$

PROOF. We add n-tuples of complex numbers entry by entry, and we multiply such an n-tuple by a complex scalar by multiplying each entry of the n-tuple by that scalar. For any n-tuples of complex numbers $a = (a_1, \dots, a_n)$ and $b = (b_1, \dots, b_n)$, define $|a| = \big(\sum_{k=1}^{n} |a_k|^2\big)^{1/2}$, $|b| = \big(\sum_{k=1}^{n} |b_k|^2\big)^{1/2}$, and $(a, b) = \sum_{k=1}^{n} a_k \overline{b_k}$.

The Schwarz inequality says that $0 \le 0$ if $b = (0, \dots, 0)$, and thus we may assume that b is something else. In this case, $|b| \ne 0$. Then

$$\begin{aligned} 0 &\le \big|a - |b|^{-2}(a, b)b\big|^2 = \big(a - |b|^{-2}(a, b)b,\ a - |b|^{-2}(a, b)b\big) \\ &= |a|^2 - 2|b|^{-2}|(a, b)|^2 + |b|^{-4}|(a, b)|^2|b|^2 = |a|^2 - |b|^{-2}|(a, b)|^2, \end{aligned}$$

and the asserted inequality follows.

A6. Equivalence Relations

An **equivalence relation** on a set S is a relation between S and itself, i.e., is a subset of $S \times S$, satisfying three properties. We define the expression $a \simeq b$, written "a is equivalent to b," to mean that the ordered pair (a, b) is a member of the relation, and we say that "$\simeq$" is the equivalence relation. The properties are

(i) $a \simeq a$ for all a in S, i.e., $\simeq$ is **reflexive**,
(ii) $a \simeq b$ implies $b \simeq a$ if a and b are in S, i.e., $\simeq$ is **symmetric**.
(iii) $a \simeq b$ and $b \simeq c$ together imply $a \simeq c$ if a, b, and c are in S, i.e., $\simeq$ is **transitive**.

An example occurs with S equal to the set $\mathbb{Z}$ of integers with $a \simeq b$ meaning that the difference $a - b$ is even. The properties hold because (i) 0 is even, (ii) the negative of an even integer is even, and (iii) the sum of two even integers is even.

There is one fundamental result about abstract equivalence relations. The **equivalence class** of a, written $[a]$ for now, is the set of all members b of S such that $a \simeq b$.

Proposition. If $\simeq$ is an equivalence relation on a set S, then any two equivalence classes are disjoint or equal, and S is the union of all the equivalence classes.

PROOF. Let $[a]$ and $[b]$ be the equivalence classes of members a and b of S. If $[a] \cap [b] \neq \varnothing$, choose c in the intersection. Then $a \simeq c$ and $b \simeq c$. By (ii), $c \simeq b$, and then by (iii), $a \simeq b$. If d is any member of $[b]$, then $b \simeq d$. From (iii), $a \simeq b$ and $b \simeq d$ together imply $a \simeq d$. Thus $[b] \subseteq [a]$. Reversing the roles of a and b, we see that $[a] \subseteq [b]$ also, whence $[a] = [b]$. This proves the first conclusion. The second conclusion follows from (i), which ensures that a is in $[a]$, hence that every member of S lies in some equivalence class.

EXAMPLE. With the equivalence relation on $\mathbb{Z}$ that $a \simeq b$ if $a - b$ is even, there are two equivalence classes—the subset of even integers and the subset of odd integers.

The first two examples of equivalence relations in this book arise in Chapter II. The first example, which is in Section II.2 and concerns a passage from "pseudometric spaces" to "metric spaces," yields equivalence classes exactly as above. The second example, which is in Section II.3, is a relation "is homeomorphic to" and implicitly is defined on the class of all metric spaces. This class is not a set, and Section A1 of this appendix suggested avoiding using classes that are not sets in order to avoid the logical paradoxes mentioned at the beginning of the appendix. There is not much problem with using general classes in this particular situation, but there is a simple approach in this situation for eliminating classes that are not sets and thereby following the suggestion of Section A1 without making an exception. The approach is to work with any subclass of metric spaces that is a set. The equivalence relation is well defined on the set of metric spaces in question, and the proposition yields equivalence classes within that set. This set can be an arbitrary subclass of the class of all metric spaces that happens to be a set, and the practical effect is the same as if the equivalence relation had been defined on the class of all metric spaces.

A7. Linear Transformations, Matrices, and Determinants

A certain amount of linear algebra, done with real or complex scalars, is taken as known. The topics of vectors, vector spaces, operations on matrices, row reduction of matrices, spanning, linear independence, bases, and dimension will not be reviewed here. This section will concentrate on the correspondence between linear transformations and matrices in the finite-dimensional case, and on the elementary properties of determinants. So as to be able to handle real and complex scalars simultaneously, we denote by $\mathbb{F}$ either $\mathbb{R}$ or $\mathbb{C}$.

The linear transformations in question will be functions with domain $\mathbb{F}^n$ and range $\mathbb{F}^m$. As is emphasized for the case $\mathbb{F} = \mathbb{R}$ in Section II.1, the members of these spaces are to be regarded as column vectors with entries in $\mathbb{F}$ even if, in order to save space, one occasionally writes them horizontally with commas between entries. This is an important convention, since it makes matrix operations and operations with linear transformations correspond to each other in the same order without the need to transpose any matrix. The standard bases for $\mathbb{F}^n$ and $\mathbb{F}^m$ are often denoted by $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_m\}$, respectively, in this book, where

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

are n-entry column vectors and

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad u_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

are m-entry column vectors.

A function $T : \mathbb{F}^n \to \mathbb{F}^m$ is a **linear function** if it satisfies $T(x + y) = T(x) + T(y)$ and $T(cx) = cT(x)$ for all x and y in $\mathbb{F}^n$ and all elements c of $\mathbb{F}$. The terms "linear transformation" and "linear map" are used also.

An example is obtained from any m-by-n matrix A with entries in $\mathbb{F}$, namely $T(x) = Ax$, the right side being a matrix product. The size of A needs emphasis: the number of rows equals the dimension of the range, and the number of columns equals the dimension of the domain.

Conversely if $T : \mathbb{F}^n \to \mathbb{F}^m$ is a linear function, then there is a unique such matrix A such that $T(x) = Ax$ for all x in $\mathbb{F}^n$: the j^{th} column of A is $T(e_j)$ for $1 \le j \le n$. For example, if $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the rotation about the origin counterclockwise through an angle θ, then $T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ and $T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$. Consequently $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Sometimes it is necessary to have a notation for the entries of a matrix A, and this text uses A_{ij} to indicate the entry of A in the i^{th} row and j^{th} column. If a matrix is defined entry by entry, the entries being M_{ij}, the text will occasionally refer to the whole matrix as $[M_{ij}]$. This convention is especially handy if M_{ij} is given by some nontrivial expression like $\partial u_i / \partial x_j$ that involves i and j.

We can give a tidy formula for the correspondence $T \leftrightarrow A$ if we define a dot product in $\mathbb{F}^m$ by

$$(a_1, \ldots, a_m) \cdot (b_1, \ldots, b_m) = a_1 b_1 + \cdots + a_m b_m$$

with no complex conjugations involved. The correspondence of a linear function T in $L(\mathbb{F}^n, \mathbb{F}^m)$ to a matrix A with entries in $\mathbb{F}$ is then given by

$$A_{ij} = T(e_j) \cdot u_i.$$

The correspondence $T \leftrightarrow A$ of linear functions to matrices carries certain vector spaces associated to T to vector spaces associated with A. The **kernel** of T, namely the set of vectors x with $T(x) = 0$, corresponds to the **null space** of A, the set of column vectors with $Ax = 0$. The **image** of T, as defined in Section A1, corresponds to the column space of A, the linear span of the columns of A. The method of row reduction of matrices shows that

$$\#\{\text{columns of } A\} = \dim(\text{null space of } A) + \dim(\text{span of rows of} A),$$

while a little argument with bases shows that

$$\dim(\text{domain of } T) = \dim(\text{kernel of } T) + \dim(\text{image of } T).$$

In these two equations the left sides are equal, and the first terms on the two right sides are equal. Therefore the second terms on the two right sides are equal, and we obtain

$$\dim(\text{span of rows of} A) = \dim(\text{span of columns of} A).$$

The common value of the two sides of this equation is called the **rank** of A or of T.

Under this correspondence of linear functions between column-vector spaces with matrices of the appropriate size, composition of linear functions corresponds to matrix product in the same written order. In other words, suppose that $T : \mathbb{F}^n \to \mathbb{F}^m$ corresponds to A of size m-by-n and that $U : \mathbb{F}^m \to \mathbb{F}^k$ corresponds to B of size k-by-m. Then $U \circ T : \mathbb{F}^n \to \mathbb{F}^k$ corresponds to BA of size k-by-n.

The **determinant** function $A \mapsto \det A$ has domain the set of all square matrices over $\mathbb{F}$ and has range $\mathbb{F}$. It is uniquely defined by the three properties

(i) $\det A$ is linear in each row of A if the other rows are held fixed,
(ii) $\det A = 0$ if two rows of A are equal,
(iii) $\det I = 1$ if I denotes the identity matrix of any size.

These properties enable one to calculate $\det A$ by row reducing the matrix A. Specifically replacement of a row by the sum of it and a multiple of another row leaves $\det A$ unchanged, multiplication of a row by a constant to make the diagonal entry be one means pulling out the diagonal entry as a scalar factor multiplying the determinant, and interchanging two rows multiplies the determinant by -1. After the row reduction is complete for a square matrix, either the reduced row-echelon form is the identity matrix and (iii) says that the determinant is 1 or else the reduced row-echelon form has a row of 0's, and (i) and (ii) imply that the determinant is 0.

The determinant function has the following additional properties, which may be regarded as consequences of (i), (ii), and (iii) above:

(iv) $\det A \neq 0$ if and only if A is invertible,
(v) $\det A = \det A^{\mathrm{tr}}$, where A^{tr} is the transpose of A,
(vi) $\det(AB) = (\det A)(\det B)$,
(vii) $\det A = \sum_\sigma (\operatorname{sgn}\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)}$ if A is n-by-n with entries $A_{i,j}$; the sum is taken over all permutations σ of $\{1, \dots, n\}$, with $\operatorname{sgn}\sigma$ denoting the sign of σ,
(viii) **(expansion by cofactors)** for $n > 1$ if $\widehat{A}_{ij}$ denotes the $(n-1)$-by-$(n-1)$ matrix obtained by deleting the i^{th} row and j^{th} column from the n-by-n matrix A, then $\det A = \sum_{j=1}^n (-1)^{i+j} A_{ij} \det \widehat{A}_{ij}$ for all i and $\det A = \sum_{i=1}^n (-1)^{i+j} A_{ij} \det \widehat{A}_{ij}$ for all j,
(ix) **(Cramer's rule)** if $\det A \neq 0$, if v is in $\mathbb{R}^n$, and if A_j denotes the matrix obtained by replacing the j^{th} column of A by v, then the j^{th} entry of the unique solution $x \in \mathbb{R}^n$ of $Ax = v$ is $x_j = \det A_j \big/ \det A$.

A8. Factorization and Roots of Polynomials

The first objective of this section is to prove unique factorization of real and complex polynomials. Let $\mathbb{F}$ denote either the reals $\mathbb{R}$ or the complex numbers $\mathbb{C}$.

We work with polynomials with coefficients in $\mathbb{F}$. These are expressions $P(X) = a_n X^n + \cdots + a_1 X + a_0$ with $a_n, \dots, a_1, a_0$ in $\mathbb{F}$. Although it is tempting to think of $P(X)$ as a function with independent variable X, it is better to identify P with the sequence $(a_0, a_1, \dots, a_n, 0, 0, \dots)$ of coefficients. For this setting, a **polynomial** (in one "indeterminate") may be defined as a sequence of members of $\mathbb{F}$ such that all terms of the sequence are 0 from some point on. The indexing of the sequence is to begin with 0. Addition, scalar multiplication, and polynomial multiplication are then defined in the expected way so as to match the operations on functions. The usual associative, commutative, and distributive laws are then valid.

Nevertheless, it is still convenient to use the notation X in writing explicit polynomials. If r is in $\mathbb{F}$, we can evaluate $P(X) = a_n X^n + \cdots + a_1 X + a_0$ at r, and the result is the number $P(r) = a_n r^n + \cdots + a_1 r + a_0$. We say that r is a **root** of P if $P(r) = 0$. The **degree** of a polynomial P, denoted by $\deg P$, is the largest integer n such that the coefficient of X^n is nonzero; the notion of "degree" is left undefined for the 0 polynomial, i.e., the polynomial all of whose coefficients are 0. A **factor** of a polynomial $A(X)$ is a polynomial $B(X)$ such that $A(X) = B(X)Q(X)$ for some polynomial $Q(X)$; we say also that $B(X)$ and $Q(X)$ **divide** $A(X)$. In this case, if B and Q are not 0, then A is not 0 and $\deg A = \deg B + \deg Q$.

Division Algorithm. If $A(X)$ and $B(X)$ are polynomials with coefficients in $\mathbb{F}$ and if $B(X)$ is not the 0 polynomial, then there exist unique polynomials $Q(X)$ and $R(X)$ such that

(a) $A(X) = B(X)Q(X) + R(X)$ and
(b) either $R(X)$ is the 0 polynomial or $\deg R < \deg B$.

REMARK. This result codifies the usual method of dividing polynomials in high-school algebra. That method writes $A(X)/B(X) = Q(X) + R(X)/B(X)$, and then one obtains the above result by multiplying by $B(X)$. The polynomial Q is the quotient in the division, and $R(X)$ is the remainder.

PROOF OF UNIQUENESS. If $A = BQ_1 + R_1$ also, then $B(Q - Q_1) = R_1 - R$. Without loss of generality, $R_1 - R$ is not the 0 polynomial since otherwise $Q - Q_1 = 0$ also. Then

$$\deg B + \deg(Q - Q_1) = \deg(R_1 - R) \leq \max\{\deg R, \deg R_1\} < \deg B,$$

and we have a contradiction.

PROOF OF EXISTENCE. If $A = 0$ or $\deg A < \deg B$, we take $Q = 0$ and $R = A$, and we are done. Otherwise we induct on $\deg A$. Assume the result for degree $\leq n - 1$, and let $\deg A = n$. Write $A = a_n X^n + A_1$ with $A_1 = 0$ or $\deg A_1 < \deg A$. Let $B = b_k X^k + B_1$ with $B_1 = 0$ or $\deg B_1 < \deg B$. Put $Q_1 = a_n b_k^{-1} X^{n-k}$. Then

$$A - BQ_1 = a_n X^n + A_1 - a_n X^n - a_n b_k^{-1} X^{n-k} B_1 = A_1 - a_n b_k^{-1} X^{n-k} B_1$$

with the right side equal to 0 or of degree $< \deg A$. Then the right side, by induction, is of the form $BQ_2 + R$, and $A = B(Q_1 + Q_2) + R$ is the required decomposition.

Corollary 1 (Factor Theorem). If r is in $\mathbb{F}$ and P is a polynomial, then $X - r$ divides P if and only if $P(r) = 0$.

PROOF. If $P = (X - r)Q$, then $P(r) = (r - r)Q(r) = 0$. Conversely let $P(r) = 0$. Taking $B(X) = X - r$ in the Division Algorithm, we obtain $P = (X - r) + R$ with $R = 0$ or $\deg R < \deg(X - r) = 1$. In either event we have $0 = P(r) = (r - r)Q(r) + R(r)$, and thus $R(r) = 0$. Of the two choices, we must have $R = 0$, and then $P = (X - r)Q$.

Proposition. If P is a nonzero polynomial with coefficients in $\mathbb{F}$ and if $\deg P = n$, then P has at most n distinct roots.

PROOF. Let $r_1, \ldots, r_{n+1}$ be distinct roots of $P(X)$. By the Factor Theorem, $X - r_1$ is a factor of $P(X)$. We prove inductively on k that the product $(X - r_1)(X - r_2)\cdots(X - r_k)$ is a factor of $P(X)$. Assume that this assertion holds for k, so that $P(X) = (X - r_1)\cdots(X - r_k)Q(X)$ and

$$0 = P(r_{k+1}) = (r_{k+1} - r_1)\cdots(r_{k+1} - r_k)Q(r_{k+1}).$$

Since the r_j's are distinct, we must have $Q(r_{k+1}) = 0$. By the Factor Theorem, we can write $Q(X) = (X - r_{k+1})R(X)$ for some polynomial $R(X)$. Substitution gives $P(X) = (X-r_1)\cdots(X-r_k)(X-r_{k+1})R(X)$, and $(X-r_1)\cdots(X-r_{k+1})$ is exhibited as a factor of $P(X)$. This completes the induction. Consequently

$$P(X) = (X - r_1)\cdots(X - r_{n+1})S(X)$$

for some polynomial $S(X)$. Comparing the degrees of the two sides, we find that $\deg S = -1$, and we have a contradiction.

A **greatest common divisor** of polynomials A and B with $B \neq 0$ is any polynomial D of maximum degree such that D divides A and D divides B. The **Euclidean algorithm** is the iterative process that makes use of the Division Algorithm in the form

$$\begin{aligned}
A &= BQ_1 + R_1, &\qquad R_1 &= 0 \text{ or } \deg R_1 < \deg B,\\
B &= R_1Q_2 + R_2, & R_2 &= 0 \text{ or } \deg R_2 < \deg R_1,\\
R_1 &= R_2Q_3 + R_3, & R_3 &= 0 \text{ or } \deg R_3 < \deg R_2,\\
&\ \vdots\\
R_{n-2} &= R_{n-1}Q_n + R_n, & R_n &= 0 \text{ or } \deg R_n < \deg R_{n-1},\\
R_{n-1} &= R_nQ_{n+1}.
\end{aligned}$$

In the above computation the integer n is defined by the conditions that $R_n \neq 0$ and that $R_{n+1} = 0$. Such an n must exist since $\deg B > \deg R_1 > \cdots \geq 0$.

Theorem. Let A and B be polynomials with coefficients in $\mathbb{F}$ and with $B \neq 0$, and let $R_1, \ldots, R_n$ be the remainders generated by the Euclidean algorithm when applied to A and B. Then

(a) R_n is a greatest common divisor of A and B,
(b) the greatest common divisor D of A and B is unique up to scalar multiplication,
(c) any D_1 that divides both A and B necessarily divides D,
(d) there exist polynomials P and Q with $AP + BQ = D$.

PROOF. Let D_1 divide A and B. From $A = BQ_1 + R_1$, we see that D_1 divides R_1. From $B = R_1Q_2 + R_2$, we see that D_1 divides R_2. Continuing in this way through $R_{n-2} = R_{n-1}Q_n + R_n$, we see that D_1 divides R_n. In particular any greatest common divisor D of A and B divides R_n and therefore has $\deg D \leq \deg R_n$. In the reverse direction, $R_{n-1} = R_nQ_{n+1}$ shows that R_n divides R_{n-1}. From $R_{n-2} = R_{n-1}Q_n + R_n$, we see that R_n divides R_{n-2}. Continuing in this way through $B = R_1Q_2 + R_2$, we see that R_n divides B. Finally $A = BQ_1 + R_1$ shows that R_n divides A and B. Thus R_n is a divisor of both A and B, and we have seen that its degree is maximal. This proves (a).

If D is a greatest common divisor of A and B, it follows that D divides R_n and $\deg D = \deg R_n$. This proves (b). We have seen that any D_1 that divides A and B necessarily divides R_n, and then (c) follows from the uniqueness of the greatest common divisor up to scalar multiplication.

Put $R_{n+1} = 0$, $R_0 = B$, and $R_{-1} = A$. We prove by induction downward that there are polynomials S_k and T_k such that $R_kS_k + R_{k+1}T_k = D$. The base case of the induction is $k = n$, where we have $R_n1 + R_{n+1}0 = D$. Suppose that $R_kS_k + R_{k+1}T_k = D$ with $k \geq 0$. We rewrite $R_{k-1} = R_kQ_{k+1} + R_{k+1}$ as $R_{k+1} = R_{k-1} - R_kQ_{k+1}$ and substitute to obtain

$$D = R_kS_k + R_{k+1}T_k = R_kS_k + R_{k-1}T_k - R_kQ_{k+1}.$$

In other words, we can take $S_{k-1} = T_k$ and $T_k = S_k - Q_{k+1}$, and our inductive assertion is proved for $k - 1$. The assertion for -1 proves (d).

A nonzero polynomial P with coefficients in $\mathbb{F}$ is **prime** if the only factors of P are the scalar multiples of 1 and the scalar multiples of P.

Lemma. If A and B are nonzero polynomials with coefficients in $\mathbb{F}$ and if P is a prime polynomial such that P divides AB, then P divides A or P divides B.

PROOF. Suppose that P does not divide A. Then 1 is a greatest common divisor of A and P, and part (d) of the above theorem produces polynomials S and T such that $AS + PT = 1$. Multiplication by B gives $ABS + PTB = B$. Then P divides ABS because it divides AB, and P divides PTB because it divides P. Hence P divides B.

Theorem (unique factorization). Every polynomial of degree ≥ 1 with coefficients in $\mathbb{F}$ is a product of primes. This factorization is unique up to order and to scalar multiplication of the prime factors.

PROOF. If A is given and is not prime, decompose $A = BC$ with $\deg B < \deg A$ and $\deg C < \deg A$. For each factor that is not prime, write the factor as the product of two polynomials of lower degree. This process, when continued in this fashion, must stop since the degrees strictly decrease with any factorization. This proves existence.

For uniqueness, assume the contrary and choose $m \geq 1$ as small as possible so that some polynomial has two distinct factorizations $P_1 \cdots P_m = Q_1 \cdots Q_n$ into primes, apart from order and scalar factors. Adjusting scalar multiples, we may assume that each P_j and Q_k has leading coefficient 1 and that there is a global coefficient multiplying each side. These global coefficients must be equal, being the coefficients of the largest power of X on each side. Thus we may cancel them and assume that each P_j and Q_k has leading coefficient 1. By the lemma, the fact that Q_1 is prime means that Q_1 must divide one of $P_1, \dots, P_m$. Reordering the factors, we may assume that Q_1 divides P_1. Since P_1 is prime, P_1 is a scalar multiple of Q_1. Since P_1 and Q_1 both have leading coefficient 1, $P_1 = Q_1$. Then we can cancel P_1 and Q_1 from both of our factorizations, obtaining distinct factorizations with fewer than m factors on one side. By the minimality of m, either we have arrived at a contradiction or we now have the polynomial 1 left on one side. Then the other side is 1, and the two sides match.

If $\mathbb{F}$ is $\mathbb{R}$, then $X^2 + 1$ is prime. But $X^2 + 1$ is not prime when $\mathbb{F} = \mathbb{C}$ since $X^2 + 1 = (X + i)(X - i)$. The Fundamental Theorem of Algebra, stated below, implies that every prime polynomial over $\mathbb{C}$ is of degree 1. It is possible to prove the Fundamental Theorem of Algebra within complex analysis as a consequence of Liouville's Theorem or within modern algebra as a consequence of Galois theory and the Sylow theorems. This text gives a proof of the result in Section II.7 using the Heine–Borel Theorem and other facts about compactness.

Fundamental Theorem of Algebra. Any polynomial with coefficients in $\mathbb{C}$ and with degree ≥ 1 has at least one root.

Corollary. Let P be a nonzero polynomial of degree n with coefficients in $\mathbb{C}$, and let $r_1, \dots, r_k$ be the roots. Then there exist unique integers $m_j > 0$ such that $P(X)$ is a multiple of $\prod_{j=1}^{k} (X - r_j)^{m_j}$. The numbers m_j have $\sum_{j=1}^{k} m_j = n$.

PROOF. We may assume that $\deg P > 0$. We apply unique factorization to $P(X)$. It follows from the Fundamental Theorem of Algebra and the Factor Theorem that each prime polynomial with coefficients in $\mathbb{C}$ has degree 1. Thus the unique factorization of $P(X)$ has to be of the form $c \prod_{l=1}^{n} (X - z_l)$ for some complex numbers that are unique up to order. The z_l's are roots, and every root is a z_l, by the Factor Theorem. Grouping like factors proves the desired factorization and its uniqueness. The numbers m_j have $\sum_{j=1}^{k} m_j = n$ by a count of degrees.

The integers m_j in the corollary are called the **multiplicities** of the roots of the polynomial $P(X)$.

A9. Partial Orderings and Zorn's Lemma

A **partial ordering** on a set S is a relation between S and itself, i.e., a subset of $S \times S$, satisfying two properties. We define the expression $a \leq b$ to mean that the ordered pair (a, b) is a member of the relation, and we say that "$\leq$" is the partial ordering. The properties are

(i) $a \leq a$ for all a in S, i.e., $\leq$ is **reflexive**,

(ii) $a \leq b$ and $b \leq c$ together imply $a \leq c$ whenever a, b, and c are in S, i.e., $\leq$ is **transitive**.

An example of such an S is any set of subsets of a set X, with $\leq$ taken to be inclusion $\subseteq$. This particular partial ordering has a third property of interest, namely

(iii) $a \leq b$ and $b \leq a$ with a and b in S imply $a = b$.

However, the validity of (iii) has no bearing on Zorn's Lemma below. A partial ordering is said to be a **total ordering** or **simple ordering** if (iii) holds and also

(iv) any a and b in S have either $a \leq b$ or $b \leq a$.

For the sake of a result to be proved at the end of the section, let us interpolate one further definition: a totally ordered set is said to be **well ordered** if every nonempty subset has a least element, i.e., if each nonempty subset contains an element a such that $a \leq b$ for all b in the subset.

A **chain** in a partially ordered set S is a totally ordered subset. An **upper bound** for a chain T is an element u in S such that $c \leq u$ for all c in T. A **maximal element** in S is an element m such that $m \leq a$ for some a in S implies $a \leq m$. (If (iii) holds, we can then conclude that $m = a$.)

Zorn's Lemma. If S is a nonempty partially ordered set in which every chain has an upper bound, then S has a maximal element.

REMARKS. Zorn's Lemma will be proved below using the Axiom of Choice, which was stated in Section A1. It is an easy exercise to see, conversely, that Zorn's Lemma implies the Axiom of Choice. It is customary with many mathematical writers to mention Zorn's Lemma each time it is invoked, even though most writers nowadays do not ordinarily acknowledge uses of the Axiom of Choice. Before coming to the proof, we give an example of how Zorn's Lemma is used.

EXAMPLE. Zorn's Lemma gives a quick proof that any real vector space V has a basis. In fact, let S be the set of all linearly independent subsets of V, and order S by inclusion upward as in the example above of a partial ordering. The set S is nonempty because $\varnothing$ is a linearly independent subset of V. Let T be a chain in S, and let u be the union of the members of T. If t is in T, we certainly have $t \subseteq u$. Let us see that u is linearly independent. For u to be dependent would mean that there are vectors $x_1, \dots, x_n$ in u with $r_1x_1 + \cdots + r_nx_n = 0$ for some system of real numbers not all 0. Let x_j be in the member t_j of the chain T. Since $t_1 \subseteq t_2$ or $t_2 \subseteq t_1$, x_1 and x_2 are both in t_1 or both in t_2. To keep the notation neutral, say they are both in t_2'. Since $t_2' \subseteq t_3$ or $t_3 \subseteq t_2'$, all of x_1, x_2, x_3 are in t_2' or they are all in t_3. Say they are both in t_3'. Continuing in this way, we arrive at one of the sets $t_1, \dots, t_n$, say t_n', such that all of $x_1, \dots, x_n$ are all in t_n'. The members of t_n' are linearly independent by assumption, and we obtain the contradiction $r_1 = \cdots = r_n = 0$. We conclude that the chain T has an upper bound in S. By Zorn's Lemma, S has a maximal element, say m. If m is not a basis, it fails to span. If a vector x is not in its span, it is routine to see that $m \cup \{x\}$ is linearly independent and properly contains m, in contradiction to the maximality of m. We conclude that m is a basis.

We now begin the proof of Zorn's Lemma. If T is a chain in a partially ordered set S, then an upper bound u_0 for T is a **least upper bound** for T if $u_0 \leq u$ for all upper bounds of T. If (iii) holds in S, then there can be at most one least upper bound for T. In fact, if u_0 and u_0' are least upper bounds, then $u_0 \leq u_0'$ since u_0 is a least upper bound, and $u_0' \leq u_0$ since u_0' is a least upper bound; by (iii), $u_0 = u_0'$.

Lemma. Let X be a nonempty partially ordered set such that (iii) holds, and write $\leq$ for the partial ordering. Suppose that X has the additional property that each nonempty chain in X has a least upper bound in X. If $f : X \to X$ is a function such that $x \leq f(x)$ for all x in X, then there exists an x_0 in X with $f(x_0) = x_0$.

PROOF. A nonempty subset E of X will be called *admissible* for purposes of this proof if $f(E) \subseteq E$ and if the least upper bound of each nonempty chain in E, which exists in X by assumption, actually lies in E. By assumption, X is an admissible subset of X. If x is in X, then the intersection of admissible subsets of X containing x is admissible. Thus the intersection A_x of all admissible subsets containing x is an admissible subset containing x. The set of all y in X with $x \leq y$ is an admissible subset of X containing x, and it follows that $x \leq y$ for all y in A_x.

By hypothesis, X is nonempty. Fix an element a in X, and let $A = A_a$. The main step is to prove that A is a chain. Once that is established, we argue as follows: Since A is a chain, its least upper bound x_0 lies in X, and since A is an admissible subset, x_0 lies in A. By admissibility, $f(A) \subseteq A$. Hence $f(x_0)$ is in A. Since x_0 is an upper bound of A, $f(x_0) \leq x_0$. On the other hand, $x_0 \leq f(x_0)$ by the assumed property of f. Therefore $f(x_0) = x_0$ by (iii).

To prove that A is a chain, consider the subset C of members x of A with the property that there is a nonempty chain C_x in A containing a and x such that

- $a \le y \le x$ for all y in C_x,
- $f(C_x - \{x\}) \subseteq C_x$, and
- the least upper bound of any nonempty subchain of C_x is in C_x.

The element a is in C because we can take $C_a = \{a\}$. If x is in C, so that C_x exists, let us use the bulleted properties to see that

$$A = A_x \cup C_x. \tag{$*$}$$

We have $A \supseteq C_x$ by definition; also $A \cap A_x$ is an admissible set containing x and hence containing A_x, and thus $A \supseteq A_x$. Therefore $A \supseteq A_x \cup C_x$. For the reverse inclusion it is enough to prove that $A_x \cup C_x$ is an admissible subset of X containing a. The element a is in C_x and thus is in $A_x \cup C_x$. For the admissibility we have to show that $f(A_x \cup C_x) \subseteq A_x \cup C_x$ and that the least upper bound of any nonempty chain in $A_x \cup C_x$ lies in $A_x \cup C_x$. Since x lies in A_x, $A_x \cup C_x = A_x \cup (C_x - \{x\})$ and $f(A_x \cup C_x) = f(A_x) \cup f(C_x - \{x\}) \subseteq A_x \cup C_x$, the inclusion following from the admissibility of A_x and the second bulleted property of C_x.

To complete the proof of $(*)$, take a nonempty chain in $A_x \cup C_x$, and let u be its least upper bound in X; it is enough to show that u is in $A_x \cup C_x$. The element u is necessarily in A since A is admissible. Observe that

$$y \le x \quad \text{and} \quad x \le z \qquad \text{whenever } y \text{ is in } C_x \text{ and } z \text{ is in } A_x. \tag{$**$}$$

If the chain has at least one member in A_x, then $(**)$ implies that $x \le u$, and hence the set of members of the chain that lie in A_x forms a nonempty chain in A_x with least upper bound u. Since A_x is admissible, u is in A_x. Otherwise the chain has all its members in C_x, and then u is in C_x by the third bulleted property of C_x.

This completes the proof of $(*)$. Although we do not need the fact, let us observe that combining $(**)$ and $(*)$ yields $A_x \cap C_x = \{x\}$ whenever C_x exists. Thus $C_x = (A - A_x) \cup \{x\}$ if C_x exists. In particular the defining properties of C_x determine C_x completely.

Recalling that C is the subset of members of A such that C_x exists, we shall show that C is an admissible set containing a. If we can do so, then it follows that $C \supseteq A_a = A$ and hence $C = A$. This fact, in combination with $(*)$ and $(**)$, proves that A is a chain: if x and y are in A, $(*)$ shows that y is in A_x or C_x, and $(**)$ shows that $x \le y$ in the first case and $y \le x$ in the second case.

Thus the proof will be complete if we show that C is admissible and contains a. We already observed that it contains a. We need to see that $f(C) \subseteq C$ and that the least upper bound of any nonempty chain of C lies in C. For the first of these conclusions, let us see that if x is in C, then $C_{f(x)}$ exists and can be taken to be $C_x \cup \{f(x)\}$. Property $(*)$ proves that $C_x \cup \{f(x)\}$ satisfies the first bulleted property of $C_{f(x)}$, and the second bulleted property follows from that same property of C_x and the fact that $x \leq f(x)$. Any nonempty chain in $C_x \cup \{f(x)\}$ either lies in C_x and has its least upper bound in C_x or else contains $f(x)$ as a member and then has $f(x)$ as its least upper bound. Thus $C_x \cup \{f(x)\}$ satisfies the third bulleted property of $C_{f(x)}$ and can be taken to be $C_{f(x)}$. Hence $f(x)$ is in C, and $f(C) \subseteq C$.

Finally take any nonempty chain $\{x_\alpha\}$ in C, and let u be its least upper bound, which is necessarily in A. Form the set $\big(\bigcup_\alpha C_{x_\alpha}\big) \cup \{u\}$. It is immediate that this set satisfies the three bulleted properties of C_u and therefore can be taken to be C_u. Hence u is in C, and C contains the least upper bound of any nonempty chain in C. Then C is indeed an admissible set containing a.

PROOF OF ZORN'S LEMMA. Let S be a partially ordered set, with partial ordering $\leq$, in which every chain has an upper bound. Let X be the partially ordered system, ordered by inclusion upward $\subseteq$, of nonempty chains[5] in S. The partially ordered system X, being given by ordinary inclusion, satisfies property (iii). A nonempty chain C in X is a nested system of chains c_α of S, and $\bigcup_\alpha c_\alpha$ is a chain in S that is a least upper bound for C. The lemma is therefore applicable to any function $f : X \to X$ such that $c \subseteq f(c)$ for all c in X. We use the lemma to produce a maximal chain in X.

Arguing by contradiction, suppose that no chain within S is maximal under inclusion. For each nonempty chain c within S, let $f(c)$ be a chain with $c \subseteq f(c)$ and $c \neq f(c)$. (This choice of $f(c)$ for each c is where we use the Axiom of Choice.) The result is a function $f : X \to X$ of the required kind, the lemma says that $f(c) = c$ for some c in X, and we arrive at a contradiction. We conclude that there is some maximal chain c_0 within S.

By assumption in Zorn's lemma, every nonempty chain within S has an upper bound. Let u_0 be an upper bound for the maximal chain c_0. If u is a member of S with $u_0 \leq u$, then $c_0 \cup \{u\}$ is a chain and maximality implies that $c_0 \cup \{u\} = c_0$. Therefore u is in c_0, and $u \leq u_0$. This is the condition that u_0 is a maximal element of S.

Corollary (Zermelo's well-ordering theorem). Every set has a well ordering.

PROOF. Let S be a set, and let $\mathcal{E}$ be the family of all pairs $(E, \leq_E)$ such that E is a subset of S and $\leq_E$ is a well-ordering of E. The family $\mathcal{E}$ is nonempty since

[5] Here a chain is simply a certain kind of subset of S, and no element of S can occur more than once in it even if (iii) fails for the partial ordering. Thus if $S = \{x, y\}$ with $x \leq y$ and $y \leq x$, then $\{x, y\}$ is in X and in fact is maximal in X.

$(\varnothing, \varnothing)$ is a member of it. We partially order $\mathcal{E}$ by a notion of "inclusion as an initial segment," saying that $(E, \leq_E) \leq (F, \leq_F)$ if

(i) $E \subseteq F$,
(ii) a and b in E with $a \leq_E b$ implies $a \leq_F b$,
(iii) a in E and b in F but not E together imply $a \leq_F b$.

In preparation for applying Zorn's Lemma, let $\mathcal{C} = \{(E_\alpha, \leq_\alpha)\}$ be a chain in $\mathcal{E}$, with the α's running through some set I. Define $E_0 = \bigcup_\alpha E_\alpha$ and define $\leq_0$ as follows: If e_1 and e_2 are in E_0, let e_1 be in E_{α_1} with α_1 in I, and let e_2 be in E_{α_2} with α_2 in I. Since $\mathcal{C}$ is a chain, we may assume without loss of generality that $(E_{\alpha_1}, \leq_{\alpha_1}) \leq (E_{\alpha_2}, \leq_{\alpha_2})$, so that $E_{\alpha_1} \subseteq E_{\alpha_2}$ in particular. Then e_1 and e_2 are both in E_{α_2} and we define $e_1 \leq_0 e_2$ if $e_1 \leq_{\alpha_2} e_2$, or $e_2 \leq_0 e_1$ if $e_2 \leq_{\alpha_2} e_1$. Because of (i) and (ii) above, the result is well defined independently of the choice of α_1 and α_2. Similar reasoning shows that $\leq_0$ is a total ordering of E_0. If we can prove that $\leq_0$ is a well ordering, then $(E_0, \leq_0)$ is evidently an upper bound in $\mathcal{E}$ for the chain $\mathcal{C}$, and Zorn's Lemma is applicable.

Now suppose that F is a nonempty subset of E_0. Pick an element of F, and let E_{α_0} be a set in the chain that contains it. Since $(E_{\alpha_0}, \leq_{\alpha_0})$ is well ordered and $F \cap E_{\alpha_0}$ is nonempty, $F \cap E_{\alpha_0}$ contains a least element f_0 relative to $\leq_{\alpha_0}$. We show that $f_0 \leq_0 f$ for all f in F. In fact, if f is given, there are two possibilities. One is that f is in E_{α_0}; in this case, the consistency of $\leq_0$ with $\leq_{\alpha_0}$ forces $f_0 \leq_0 f$. The other is that f is not in E_{α_0} but is in some E_{α_1}. Since $\mathcal{C}$ is a chain and $E_{\alpha_1} \subseteq E_{\alpha_0}$ fails, we must have $(E_{\alpha_0}, \leq_{\alpha_0}) \leq (E_{\alpha_1}, \leq_{\alpha_1})$. Then f is in E_{α_1} but not E_{α_0}, and property (iii) above says that $f_0 \leq_{\alpha_1} f$. By the consistency of the orderings, $f_0 \leq_0 f$. Hence f_0 is a least element in F, and E_0 is well ordered.

Application of Zorn's Lemma produces a maximal element $(E, \leq_E)$ of $\mathcal{E}$. If E were a proper subset of S, we could adjoin to E a member s of S not in E and define every element e of E to be $\leq s$. The result would contradict maximality. Therefore $E = S$, and S has been well ordered.

A10. Cardinality

Two sets A and B are said to have the same **cardinality**, written card $A =$ card B, if there exists a one-one function from A onto B. On any set $\mathcal{A}$ of sets, "having the same cardinality" is plainly an equivalence relation and therefore partitions $\mathcal{A}$ into disjoint equivalence classes, the sets in each class having the same cardinality. The question of what constitutes cardinality (or a "cardinal number") in its own right is one that is addressed in set theory but that we do not need to address carefully here; the idea is that each equivalence class under "having the same cardinality" has a distinguished representative, and the **cardinal number** is defined to be that representative. We write card A for the cardinal number of a set A.

Having addressed equality, we now introduce a partial ordering, saying that card $A \le$ card B if there is a one-one function from A into B. The first result below is that card $A \le$ card B and card $B \le$ card A together imply card $A =$ card B.

Proposition (Schroeder–Bernstein Theorem). If A and B are sets such that there exist one-one functions $f : A \to B$ and $g : B \to A$, then A and B have the same cardinality.

PROOF. Define the function $g^{-1} : \operatorname{image} g \to A$ by $g^{-1}(g(a)) = a$; this definition makes sense since g is one-one. Write $(g \circ f)^{(n)}$ for the composition of $g \circ f$ with itself n times, and define $(f \circ g)^{(n)}$ similarly. Define subsets A_n and A'_n of A and subsets B_n and B'_n for $n \ge 0$ by

$$\begin{aligned} A_n &= \operatorname{image}((g \circ f)^{(n)}) - \operatorname{image}((g \circ f)^{(n)} \circ g), \\ A'_n &= \operatorname{image}((g \circ f)^{(n)} \circ g) - \operatorname{image}((g \circ f)^{(n+1)}), \\ B_n &= \operatorname{image}((f \circ g)^{(n)}) - \operatorname{image}((f \circ g)^{(n)} \circ f), \\ B'_n &= \operatorname{image}((f \circ g)^{(n)} \circ f) - \operatorname{image}((f \circ g)^{(n+1)}), \end{aligned}$$

and let

$$A_\infty = \bigcap_{n=0}^\infty \operatorname{image}((g \circ f)^{(n)}) \qquad \text{and} \qquad B_\infty = \bigcap_{n=0}^\infty \operatorname{image}((f \circ g)^{(n)}).$$

Then we have

$$A = A_\infty \cup \bigcup_{n=0}^\infty A_n \cup \bigcup_{n=0}^\infty A'_n \qquad \text{and} \qquad B = B_\infty \cup \bigcup_{n=0}^\infty B_n \cup \bigcup_{n=0}^\infty B'_n,$$

with both unions disjoint.

Let us prove that f carries A_n one-one onto B'_n. If a is in A_n, then $a = (g \circ f)^{(n)}(x)$ for some $x \in A$ and a is not of the form $(g \circ f)^{(n)}(g(y))$ with $y \in B$. Applying f, we obtain $f(a) = (f \circ ((g \circ f)^{(n)})(x) = (f \circ g)^{(n)}(f(x))$, so that $f(a)$ is in the image of $((f \circ g)^{(n)} \circ f)$. Meanwhile, if $f(a)$ is in the image of $(f \circ g)^{(n+1)}$, then $f(a) = (f \circ g)^{(n+1)}(y) = f((g \circ f)^{(n)}(g(y)))$ for some $y \in B$. Since f is one-one, we can cancel the f on the outside and obtain $a = (g \circ f)^{(n)}(g(y))$, in contradiction to the fact that a is in A_n. Thus f carries A_n into B'_n, and it is certainly one-one. To see that $f(A_n)$ contains all of B'_n, let $b \in B'_n$ be given. Then $b = (f \circ g)^{(n)}(f(x))$ for some $x \in A$ and b is not of the form $(f \circ g)^{(n+1)}(y)$ with $y \in B$. Hence $b = f((g \circ f)^{(n)}(x))$, i.e., $b = f(a)$ with $a = (g \circ f)^{(n)}(x)$. If this element a were in the image of $(g \circ f)^{(n)} \circ g$, we could write $a = (g \circ f)^{(n)}(g(y))$ for some $y \in B$, and then we would have

$b = f(a) = f((g \circ f)^{(n)}(g(y))) = (f \circ g)^{(n+1)}(y)$, contradiction. Thus a is in A_n, and f carries A_n one-one onto B'_n.

Similarly g carries B_n one-one onto A'_n. Since A'_n is in the image of g, we can apply g^{-1} to it and see that g^{-1} carries A'_n one-one onto B_n.

The same kind of reasoning as above shows that f carries A_∞ one-one onto B_∞. In summary, f carries each A_n one-one onto B'_n and carries A_∞ one-one onto B_∞, while g^{-1} carries each A'_n one-one onto B_n. Then the function

$$h = \begin{cases} f & \text{on } A_\infty \text{ and each } A_n, \\ g^{-1} & \text{on each } A'_n, \end{cases}$$

carries A one-one onto B.

Next we show that any two sets A and B have comparable cardinalities in the sense that either $\operatorname{card} A \le \operatorname{card} B$ or $\operatorname{card} B \le \operatorname{card} A$.

Proposition. If A and B are two sets, then either there is a one-one function from A into B or there is a one-one function from B into A.

PROOF. Consider the set S of all one-one functions $f : E \to B$ with $E \subseteq A$, the empty function with $E = \varnothing$ being one such. Each such function is a certain subset of $A \times B$. If we order S by inclusion upward, then the union of the members of any chain is an upper bound for the chain. By Zorn's Lemma let $G : E_0 \to B$ be a maximal one-one function of this kind, and let F_0 be the image of G. If $E_0 = A$, then G is a one-one function from A into B. If $F_0 = B$, then G^{-1} is a one-one function from B into A. If neither of these things happens, then there exist $x_0 \in A - E_0$ and y_0 in $B - F_0$, and the function $\widetilde{G}$ equal to G on E_0 and having $\widetilde{G}(x_0) = y_0$ extends G and is still one-one; thus it contradicts the maximality of G.

Cantor's proof that there exist uncountable sets, done with a diagonal argument, in fact showed how to start from any set A and construct a set with strictly larger cardinality.

Proposition (Cantor). If A is a set and 2^A denotes the set of all subsets of A, then $\operatorname{card} 2^A$ is strictly larger than $\operatorname{card} A$.

PROOF. The map $A \mapsto \{A\}$ is a one-one function from A into 2^A. If we are given a one-one function $F : A \to 2^A$, let E be the set of all x in A such that x is not in $F(x)$. If $F(x_0) = E$, then $x_0 \in E$ implies $x_0 \notin F(x_0) = E$, while $x_0 \notin E$ implies $x \in F(x_0) = E$. We have a contradiction in any case, and hence E is not in the image of F. We conclude that F cannot be onto 2^A.

HINTS FOR SOLUTIONS OF PROBLEMS

Chapter I

1. Let E be a nonempty set that is bounded above. Start with a member s_1 of E. Choose if possible an s_2 in E with $s_2 - s_1 \geq 1$. Continue with $s_3 - s_2 \geq 1, s_4 - s_3 \geq 1$, etc., until this is no longer possible; the existence of an upper bound forces the process to stop at some stage. Suppose that s_k has been constructed at this stage. Define s_{k+n} inductively for $n \geq 1$ to be a member of E with $s_{k+n} - s_{k+n-1} \geq 2^{-n}$ if possible; otherwise define $s_{k+n} = s_{k+n-1}$. Then $\{s_n\}$ is bounded and monotone increasing. To complete the problem, one has only to show that $\lim_n s_n$ is the least upper bound of E.

2. Show that $x_1 \geq \sqrt{a}$ and that $\sqrt{a} \leq x_{n+1} \leq x_n$ for $n \geq 1$. Then $\lim_n x_n = c$ exists by Corollary 1.6, and c must satisfy $c = \frac{1}{2}(c^2 + a)/c$.

3. Write out a few cases and guess that the pattern is $a_{2n} = \frac{1}{2}(1 - 2^{-(n-1)})$ for $n \geq 1$ and $a_{2n+1} = 1 - 2^{-n}$ for $n \geq 0$. Prove each of these statements by induction. Since $a_{2n} \to \frac{1}{2}$ and $a_{2n+1} \to 1$ and since these two subsequences use all the terms of the sequence, the only subsequential limits of $\{a_k\}$ are $\frac{1}{2}$ and 1. Therefore $\limsup a_k = 1$ and $\liminf a_k = \frac{1}{2}$.

4. The argument without paying attention to finiteness is that $a_n + b_n \leq \sup_{r \geq k} a_r + \sup_{r \geq k} b_r$ for $n \geq k$, then that $\sup_{r \geq k}(a_r + b_r) \leq \sup_{r \geq k} a_r + \sup_{r \geq k} b_r$ for all r, and then that the limit of the sum is the sum of the limits.

5. Only (ii) converges uniformly, the reason being that $0 \leq x^n/n \leq 1/n$ and that $\lim 1/n = 0$. There is uniform convergence in (i) on $[0, 1-\epsilon]$ because $0 \leq x^n \leq (1-\epsilon)^n$, and there is uniform convergence in (iii) on $[0, 1-\epsilon]$ because the Weierstrass M test applies with $|x^k|/k \leq (1-\epsilon)^k$ and $\sum_k (1-\epsilon)^k < +\infty$.

6. The uniform convergence of $\sum_{n=0}^{\infty} a_n(x)$ follows from Corollary 1.18, and the pointwise convergence of $\sum_{n=0}^{\infty} |a_n(x)|$ follows because $(1-x)\sum_{n=0}^{\infty} x^n = 1$ for $0 \leq x < 1$ and because every $a_n(x)$ is 0 for $x = 1$. The convergence of $\sum_{n=0}^{\infty} |a_n(x)|$ cannot be uniform because the sum is discontinuous and Theorem 1.21 says that it would have to be continuous.

7. Put $g_n = f - f_n$, so that g_n is continuous and decreases pointwise to the 0 function. Let $x = x_n$ be a point where $g_n(x)$ is a maximum, and let $M_n = g_n(x_n)$. We are to prove that M_n tends to 0. Suppose it does not. If $k \geq n$, then $M_k = g_k(x_k) \geq g_k(x_n) \geq g_n(x_n) = M_n$. So M_n decreases to some $M > 0$. Passing to a

subsequence if necessary, we may assume by the Bolzano–Weierstrass Theorem that $\lim_n x_n = x'$. For $k \geq n$, we have $g_k(x_n) \geq g_n(x_n) = M_n \geq M$. Letting n tend to infinity gives $g_k(x') \geq M$ since g_k is continuous. This inequality for all k contradicts the assumption that $\lim_k g_k(x') = 0$.

8. The idea is to prove the four inequalities

$$\sum_{k=0}^{2m} (-1)^k x^{2k+1}/(2k+1)! > \sin x, \qquad \sum_{k=0}^{2m+1} (-1)^k x^{2k}/(2k)! < \cos x,$$
$$\sum_{k=0}^{2m+1} (-1)^k x^{2k+1}/(2k+1)! < \sin x, \qquad \sum_{k=0}^{2m+2} (-1)^k x^{2k}/(2k)! > \cos x$$

together by an induction. They are to be proved in order for $m = 0$, then in order for $m = 1$, and so on. In each case of the inductive step, the left side minus the right side is 0 at $x = 0$ and has derivative equal to the previous left side minus right side. The Mean Value Theorem says that each left side minus right side at $x > 0$ equals the product of x and the left side minus right side at ξ with $0 < \xi < x$. Substituting the previously proved inequality at ξ then gives the result. In other words, everything comes down to proving the first inequality, namely $x > \sin x$ for $x > 0$. Arguing in the same style, we have $x - \sin x = 1 - \cos \xi$ with $0 < \xi < x$. So at least $x - \sin x \geq 0$. For $0 < x \leq \pi$, we actually obtain $x - \sin x > 0$. Since $\frac{d}{dx}(x - \sin x) \geq 0$, we have $x - \sin x \geq \pi - \sin \pi$ for $\pi \leq x$. Thus $x - \sin x > 0$ for all $x > 0$.

9. The thing to prove is that the remainder term $\frac{1}{n!}\int_0^x (x-t)^n f^{(n+1)}(t)\,dt$ tends to 0 for each x as n tends to ∞. If $x \geq 0$, the absolute value is $\leq (n!)^{-1} \int_0^x (x-t)^n\,dt = x^{n+1}/(n+1)!$, which tends to 0 for any fixed x. If $x \leq 0$, one argues in a similar fashion.

10. By a diagonal process we can find a subsequence $\{F_{n_k}\}$ convergent for each rational x. Let F be the resulting limit function, carrying the rationals in $[-1, 1]$ into $[0, 1]$. If r and s are rationals with $r \leq s$, then $F(r) = \lim_k F_{n_k}(r) \leq \lim_k F_{n_k}(s) = F(s)$. Thus F is nondecreasing on the rationals. For each real x with $-1 < x < 1$, define $F(x^-)$ to be the limit of $F(r)$ with r rational as r increases to 1, and define $F(x^+)$ to be the limit of $F(r)$ with r rational as r decreases to 1. Then $F(x^-) \leq F(x^+)$ for each x, and $F(x^+) \leq F(y^-)$ if $x < y$. For each $N > 0$, it follows that there can be only finitely many x's for which $F(x^+) - F(x^-) \geq 1/N$, and hence there can be at most countably many x's for which $F(x^-) \neq F(x^+)$. Let this exceptional set be denoted by C. For x not in C, define $F(x) = F(x^+) = F(x^-)$.

For x not in C, let us show that $\lim_k F_{n_k}(x)$ exists and equals $F(x)$. If $r < x$ is rational, we have $F(r) = \liminf_k F_{n_k}(r) \leq \liminf_k F_{n_k}(x)$; taking the supremum over r gives $F(x) = F(x^-) \leq \liminf_k F_{n_k}(x)$. Arguing similarly with s rational and $x < s$, we have $\limsup_k F_{n_k}(x) \leq \limsup_k F_{n_k}(s) = F(s)$, and hence $\limsup_k F_{n_k}(x) \leq F(x^+) = F(x)$. Combining these two conclusions, we see that $\liminf_k F_{n_k}(x) = \limsup_k F_{n_k}(x)$ and that the common value of these limits is $F(x)$.

Thus $\{F_{n_k}(x)\}$ converges except possibly for x in C. At each point of C, the sequence is bounded. Since C is countable, another use of a diagonal process produces a subsequence of F_{n_k} that converges at every point of C, hence at every point of $[-1, 1]$.

11. If $|x| > 1/\limsup \sqrt[n]{|a_n|}$, then $\sqrt[n]{|a_n|} \geq 1/|x|$ for infinitely many n. Thus $|a_n x^n| \geq 1$ for infinitely many n, and the terms of the series do not tend to 0. Hence the series cannot converge. In the reverse direction we want to see that the inequality $|x| < 1/\limsup \sqrt[n]{|a_n|}$ implies convergence of the series. We rewrite this as $\limsup \sqrt[n]{|a_n|} < 1/|x|$. Choose a number r with $\limsup \sqrt[n]{|a_n|} < r < 1/|x|$. Then $\sqrt[n]{|a_n|} \leq r$ for all sufficiently large n, $\sqrt[n]{|a_n|}\,|x| \leq r|x| < 1$ for all n sufficiently large, and $|a_n x^n| \leq (r|x|)^n$ for all n sufficiently large. Thus $\sum |a_n x^n|$ is dominated term-by-term (from some point on) by the geometric series $\sum s^n$, where $s = r|x|$. Since $s < 1$, the geometric series converges, and hence so does $\sum |a_n x^n|$.

12. $1/(1-x)^2 = \sum_{n=0}^{\infty}(n+1)x^n$, $\log(1-x) = -\sum_{n=1}^{\infty} x^n/n$, $1/(1+x^2) = \sum_{n=0}^{\infty}(-1)^n x^{2n}$, and $\arctan x = \sum_{n=0}^{\infty}(-1)^n x^{2n+1}/(2n+1)$. All these series have radius of convergence 1.

13. The proof of existence of $\arccos x$ uses the proposition in Section A3 of the appendix. The result of the calculation of the derivative is that $\frac{d}{dx}\arccos x = -1/\sqrt{1-x^2}$ for $|x| < 1$. Then $\arcsin x + \arccos x$ has derivative 0 on $(0, \pi/2)$ and hence is constant. The constant is evaluated by putting $x = 0$, and the result is that $\arcsin x + \arccos x = \pi/2$ on $(0, \pi/2)$.

14. The uniform version of Abel's Theorem is this: Let $\{a_n(x)\}_{n\geq 0}$ be a sequence of complex-valued functions with $\sum_{n=0}^{\infty} a_n(x)$ converging uniformly to the limit $s(x)$. Then $\lim_{r\uparrow 1}\sum_{n\geq 0} a_n(x) r^n = s(x)$ uniformly in x. The proof is just a matter of seeing that the estimates in the proof of Theorem 1.48 can be made uniform in x under the stated assumptions. The result about Cesàro sums is handled similarly.

15. Write $\cos n\theta = \frac{1}{2}(e^{in\theta} + e^{-in\theta})$ and $\sin n\theta = \frac{1}{2i}(e^{in\theta} - e^{-in\theta})$. Then $\sum_{n=1}^{N}\cos n\theta = \frac{1}{2}\sum_{n=1}^{N} e^{in\theta} + \frac{1}{2}\sum_{n=1}^{N} e^{-in\theta} = \frac{1}{2}\,\dfrac{1-e^{i(N+1)\theta}}{1-e^{i\theta}} + \frac{1}{2}\,\dfrac{1-e^{-i(N+1)\theta}}{1-e^{-i\theta}}$. Each numerator is bounded by 2, and each denominator gets close to 0 only as θ tends to a multiple of 2π. This proves the estimate for the cosines, and the estimate for the sines works in the same way.

17. For (a), the relevant result is that when all a_n are 0, $\sum_{n=1}^{\infty}|b_n|^2$ equals $\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi}|f(x)|^2\,dx$. Here $\displaystyle\sum_{n=1}^{\infty}|b_n|^2$ is $(4/\pi)^2\displaystyle\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}$, and $\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi}|f(x)|^2\,dx$ is just $\dfrac{2\pi}{\pi} = 2$. Hence $\displaystyle\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2} = \frac{\pi^2}{8}$.

18. We have $F(x)f(y) = \int_0^x f(t)f(y)\,dt = \int_0^x f(t+y)\,dt = \int_y^{x+y} f(t)\,dt = F(x+y) - F(y)$. If $F(x) \neq 0$ for some x, we can divide and use the Fundamental Theorem of Calculus to see that $f(y)$ has a continuous derivative everywhere. (If $F(x) = 0$ for all x, then differentiation gives $f(x) = 0$ for all x.) Differentiating the original

identity in x gives $f'(x)f(y) = f'(x+y)$. When $x = 0$, we obtain $f'(0)f(y) = f'(y)$. Then $\frac{d}{dy}\big(f(y)e^{-f'(0)y}\big) = f'(y)e^{-f'(0)y} + f(y)(-f'(0)e^{-f'(0)y}) = 0$, and hence $f(y)e^{-f'(0)y}$ is constant. Thus $f(y) = ae^{f'(0)y}$. In the original identity $f(x)f(y) = f(x+y)$, if we put $x = 0$ and choose y such that $f(y) \neq 0$, then we see that $f(0) = 1$. Hence $f(y) = e^{f'(0)y}$ if f is not identically 0.

19. We may assume that f is not identically 0. As in Problem 18, we have $f(0) = 1$. By continuity of f, choose x_0 such that $|f(x) - 1| \le \frac{1}{10}$ when $|x| \le |x_0|$. Then $\operatorname{Re} f(x_0) > 0$, and we can choose a unique c with $|\operatorname{Im}(cx_0)| < \pi/2$ such that $e^{cx_0} = f(x_0)$. The equation for f shows that $f\big(\frac{1}{2}x_0\big)^2 = f(x_0)$, and hence $f\big(\frac{1}{2}x_0\big)$ equals $e^{cx_0/2}$ or $-e^{cx_0/2}$. From $|f\big(\frac{1}{2}x_0\big) - 1| \le \frac{1}{10}$, we have $\operatorname{Re} f\big(\frac{1}{2}x_0\big) > 0$. Since $|\operatorname{Im}(cx_0/2)| < \pi/2$, $e^{cx_0/2}$ is the choice of square root of e^{cx_0} with positive real part, and we conclude that $f\big(\frac{1}{2}x_0\big) = e^{cx_0/2}$. Iterating this argument, we obtain $f(2^{-n}x_0) = e^{c2^{-n}x_0}$ for all $n \ge 0$. The equation for f shows that $f(kx) = f(x)^k$ for all integers $k \ge 0$, and thus $f(qx_0) = e^{cqx_0}$ for every rational q of the form $k/2^n$ with k an integer ≥ 0. From $f(x)f(-x) = f(0) = 1$, we have $f(x^{-1}) = f(x)^{-1}$, and thus $f(qx_0) = e^{cqx_0}$ for every rational number of the form $k/2^n$ with k any integer. Using continuity and passing to the limit, we obtain $f(r) = e^{cr}$ for all real r.

21. This uses the discussion at the end of Section A2 of the appendix. For $x \neq 0$, we compute that $g'(x) = (R(x)/S(x))e^{-1/x^2}$ for polynomials R and S with S not the 0 polynomial. Then $\lim_{x\to 0} g'(x) = 0$ by Problem 20, and the appendix shows that $g'(0)$ exists and equals 0.

22. Use Problem 21 and induction.

23. Since $\{s_n\}$ is convergent, it is bounded. Say $|s_n| \le K$ for all n. Let $\epsilon > 0$ be given, and choose N such that $n \ge N$ implies $|s_n - s| < \epsilon/2$. Write $t_n - s = \sum_j M_{nj}s_j - s = \sum_j M_{nj}(s_j - s)$ by (i). A second application of (i) gives

$$\begin{aligned}|t_n - s| &\le \sum_{j=0}^{N} M_{nj}(|s_j| + |s|) + \sum_{j=N+1}^{\infty} M_{nj}|s_j - s| \\ &\le 2K\sum_{j=0}^{N} M_{nj} + \sum_{j=N+1}^{\infty} M_{nj}\epsilon/2 \le 2K\sum_{j=0}^{N} M_{nj} + \epsilon/2.\end{aligned}$$

Since N is fixed, (ii) shows that $2K\sum_{j=0}^{N} M_{nj} < \epsilon/2$ for n sufficiently large. For those n, $|t_n - s| < \epsilon$.

24. For Cesàro summability the i^{th} row, for $i \ge 1$, has its first i entries equal to $1/i$ and its remaining entries equal to 0. For Abel summability the row going with r_i has j^{th} entry $(1 - r_i)(r_i)^j$ for $j \ge 0$.

25. Certainly $M_{ij} \ge 0$ for all i and j. The power series in Problem 12a shows that $\sum_j M_{ij} = 1$ for all i, and (ii) holds because $\lim_{r\uparrow 1}(j+1)r^j(1-r)^2 = (j+1)\cdot 1\cdot 0 = 0$.

26. Check that M as in the previous problem transforms the Cesàro sums into the Abel sums, and apply Problem 23.

27. This is handled by the same kind of computation as with the Fejér kernel.

28. The formula for $P_r(\theta)$ comes from summing the two geometric series for $n \geq 0$ and $n < 0$ and then adding the results. Properties (i) and (iii) are then immediate by inspection. For property (ii) we use the series expansion of $P_r(\theta)$. Theorem 1.31 allows the integration to be done term by term, and the result follows.

29. This is proved in the same way as Fejér's Theorem (Theorem 1.59).

30. Corollary 1.38 shows that $f_k'(x) = \sum_{n=0}^\infty c_{n,k} n x^{n-1}$ and that $f_k''(x) = \sum_{n=0}^\infty c_{n,k} n(n-1)x^{n-2}$ for $|x| < R$. The point is to show that $\{f_k'(x)\}$ is uniformly bounded and uniformly equicontinuous for $|x| \leq r$, and then Ascoli's Theorem produces the required subsequence. For proving the equicontinuity, it is enough to prove that $\{f_k''(x)\}$ is uniformly bounded for $|x| \leq r$.

Fix $r < R$, and choose r_1 with $r < r_1 < R$. Since $\lim f_k(x) = f(x)$ uniformly for $|x| \leq r_1$, there is an M such that $|f_k(r_1)| \leq M$ for all k. Thus $|\sum_n c_{n,k} r_1^n| \leq M$ for all k. Since $c_{n,k} \geq 0$ for all n and k, $c_{n,k} \leq M r_1^{-n}$ for all n and k. Since $r < r_1$, choose N such that $n \geq N$ implies $n(r/r_1)^{n-1} \leq 1$ and $n(n-1)(r/r_1)^{n-2} \leq 1$ for $n \geq N$. Since $c_{n,k} \geq 0$ for all n and k, $c_{n,k} n|x|^{n-1} \leq c_{n,k} n r^{n-1} \leq (c_{n,k} r_1^{n-1})(n(r/r_1)^{n-1}) \leq c_{n,k} r_1^{n-1}$ for $n \geq N$ and $|x| \leq r$. Summing on $n \geq N$ and taking Corollary 1.38 into account, we see that

$$\Big| f_k'(x) - \sum_{n=0}^{N-1} n c_{n,k} x^{n-1} \Big| \leq r_1^{-1} \Big(f_k(r_1) - \sum_{n=0}^{N-1} c_{n,k} r_1^n \Big) \leq r_1^{-1} f_k(r_1) \leq r_1^{-1} M$$

for $|x| \leq r$. Thus $|x| \leq r$ implies that $|f_k'(x)|$ is

$$\leq r_1^{-1} M + \sum_{n=0}^{N-1} n c_{n,k} |x|^{n-1} \leq r_1^{-1} M + \sum_{n=0}^{N-1} n c_{n,k} r_1^{n-1} \leq r_1^{-1} M + N(N-1) M r_1^{-1},$$

and $\{f_k'(x)\}$ is uniformly bounded for $|x| \leq r$.

A similar argument with f_k'' shows that

$$\Big| f_k''(x) - \sum_{n=0}^{N-1} n(n-1) c_{n,k} x^{n-2} \Big| \leq r_1^{-2} M,$$

and we find similarly that $\{f_k''(x)\}$ is uniformly bounded for $|x| \leq r$. This completes the proof.

31. Theorem 1.23 shows that the limit of the subsequence of first derivatives is the first derivative of the limit, the limit being differentiable. In other words, f is differentiable for $|x| < r$, and the subsequence converges to $f'(x)$ there. Since $r < R$ is arbitrary, f is differentiable for $|x| < R$. Now we can induct, replacing f and the sequence f_k in Problem 30 by f' and a subsequence of f_k' on a smaller disk, then passing to f'', and so on. The result is that f is infinitely differentiable for $|x| < R$.

32. This is proved in the same way as in Problem 9.

33. $|\frac{1}{N+k}z^k| \le r^{N+k}$ if $|z| \le r$, and $\sum_{k=0}^{\infty} r^{N+k} = r^N/(1-r)$. Thus $|\frac{1}{N}z^N + \frac{1}{N+1}z^{N+1} + \cdots|$ tends uniformly to 0 for $|z| \le r$. Since $t \mapsto \exp(t)$ is continuous at $t = 0$, the required convergence follows.

34. Corollary 1.38 shows from the behavior for z real that all c_n are 0.

35. Write

$$\exp\big(z + \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 + \cdots\big) = \big(\textstyle\prod_{k=1}^{N-1} \exp(\tfrac{1}{k}z^k)\big) \exp\big(\tfrac{1}{N}z^N + \tfrac{1}{N+1}z^{N+1} + \cdots\big).$$

Problem 33 shows that the left side is the uniform limit of $\prod_{k=1}^{N-1} \exp(\frac{1}{k}z^k)$ for $|z| \le r$ if $r < 1$. Each factor of the finite product is given by a convergent power series with nonnegative coefficients, and Theorem 1.40 shows that the finite product is given by a convergent power series with nonnegative coefficients. By Problem 32, $\exp\big(z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \cdots\big)$ is given by a convergent power series for $|z| < 1$. Hence $\exp\big(z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \cdots\big) - 1/(1-z)$ is given by a convergent power series for $|z| < 1$. For $z = x$ real with $|x| < 1$, the series expansion of Problem 12b shows that our expression is $\exp\big(-\log(1-x)\big) - 1/(1-x) = 0$. Thus our power series sums to 0 on the real axis. By Problem 34, it sums to 0 everywhere.

Chapter II

1. Let us compare $d(x, y)$ with $d(x, z) + d(z, y)$. If j contributes to $d(x, y)$, then $x_j \ne y_j$. Hence $x_j \ne z_j$ or $z_j \ne y_j$. Thus j contributes to at least one of $d(x, z)$ and $d(z, y)$. In other words, the contribution of j to $d(x, y)$ is $\le$ the contribution of j to $d(x, z) + d(z, y)$. Summing on j gives the desired result.

2. Let (X, d) be the given separable metric space, define E to be the subset of members x of X such that every neighborhood of x is uncountable, and let F be the complement of E. If x is in F, we can associate to x some open neighborhood N_x containing at most countably many elements, and N_x is entirely contained in F. As x varies in F, the sets N_x form an open cover of F. By Proposition 2.32b, some subcollection of the N_x that is at most countable covers F. The union of these sets is open and is at most countable, and it equals F.

3. Let $f(x) = 1/x$ for $0 < x \le 1$, and let $f(0) = 0$.

4. Suppose that x is in U. Since A is dense, the set $A \cap B(1/n; x)$ is nonempty for each $n \ge 1$. Let x_n be a member of it. Since U is open, $B(1/n; x)$ is contained in U if n is $\ge N$ for a suitable N. Thus x_n is in $A \cap U$ for $n \ge N$ and converges to x. By Proposition 2.22b, either $x_n = x$ infinitely often, in which case x is in $A \cap U$, or x is a limit point of $A \cap U$. In either case, $U \subset (A \cap U)^{\mathrm{cl}}$.

5. For (a), the sets E_n are compact by the Heine–Borel Theorem. Then each $E_n - U$ is compact. Their intersection is $\bigcap_{n=1}^{\infty} (E_n \cap U^c) = \big(\bigcap_{n=1}^{\infty} E_n\big) \cap U^c \subseteq U \cap U^c = \varnothing$. By Proposition 2.35 the system $\{E_n - U\}$ does not have the finite-intersection property. Thus $\bigcap_{n=1}^{N} (E_n - U) = \varnothing$ for some N. Since $E_1 \supseteq E_2 \supseteq \cdots$, we find that $E_N - U = \varnothing$. Therefore $E_N \subseteq U$.

For (b), let U be empty, and let $E_n = \mathbb{Q} \cap [\sqrt{2}, \sqrt{2} + 1/n]$.

6. In both parts of the problem, let the metrics be d_X, d_Y, d_Z. For (a), use continuity of F to choose for each (x_0, y) some $\delta_{1,y} > 0$ and $\delta_{2,y} > 0$ such that the two inequalities $d_X(x, x_0) < \delta_{1,y}$ and $d_Y(y', y) < \delta_{2,y}$ together imply $d_Z(F(x, y'), F(x_0, y)) < \epsilon/2$. As y varies, the open balls $B(\delta_{2,y}; y)$ cover Y. Since Y is compact, a finite number of them suffice to cover Y, say $B(\delta_{2,y_1}; y_1), \dots, B(\delta_{2,y_n}; y_n)$. Put $\delta_1 = \min\{\delta_{1,y_1}, \dots, \delta_{1,y_n}\}$. Suppose now that $d_X(x, x_0) < \delta_1$ and that y' is in Y. Then y' is in some $B(\delta_{2,y_j}; y_j)$. Hence we have $d_X(x, x_0) < \delta_1 \le \delta_{1,y_j}$ and $d(y', y_j) < \delta_{2,y_j}$, and we therefore obtain $d_Z(F(x, y'), F(x_0, y_j)) < \epsilon/2$. Since also $d_X(x_0, x_0) = 0$ and $d(y', y_j) < \delta_{2,y_j}$, we obtain also $d_Z(F(x_0, y'), F(x_0, y_j)) < \epsilon/2$. Combining these two results gives $d_Z(F(x, y'), F(x_0, y')) < \epsilon$.

For (b), consider $d_Z(F(x, y), F(x_0, y_0))$, and let $\epsilon > 0$ be given. By uniform convergence, choose $\delta_1 > 0$ such that $d_X(x, x_0) < \delta_1$ implies $d_Z(F(x, y), F(x_0, y)) < \epsilon/2$ for all y. Proposition 2.21 gives us continuity of $F(x_0, \cdot)$, and thus there exists $\delta_2 > 0$ such that $d_Y(y, y_0) < \delta_2$ implies $d_Z(F(x_0, y), F(x_0, y_0)) < \epsilon/2$. Then $d_X(x, x_0) < \delta_1$ and $d_Y(y, y_0) < \delta_2$ together imply $d_Z(F(x, y), F(x_0, y_0)) \le d_Z(F(x, y), F(x_0, y)) + d_Z(F(x_0, y), F(x_0, y_0)) < \epsilon/2 + \epsilon/2 = \epsilon$.

7. Let $f : (0, 1) \to \mathbb{R}$ be defined by $f(x) = 1/x$. Then the Cauchy sequence $\{1/n\}$ is carried to a sequence that is not Cauchy in $\mathbb{R}$.

8. Define inductively $f^{(0)}$ to be the identity and $f^{(k)} = f \circ f^{(k-1)}$ for $k > 0$. For existence we see inductively that $d(f^{(k)}(x), f^{(k)}(y)) \le r^k d(x, y)$ for all x and y. If $n \ge m$ and if x is arbitrary but fixed, we then have $d(f^{(n)}(x), f^{(m)}(x)) \le \sum_{k=m}^{n-1} d(f^{(k+1)}(x), f^{(k)}(x)) \le \sum_{k=m}^{n-1} r^k d(f(x), x) \le r^m d(f(x), x)/(1-r)$. Hence the sequence $\{f^{(n)}(x)\}$ is Cauchy. Let x' be its limit. Since $d(f(f^{(n)}(x)), f^{(n)}(x)) = d(f^{(n+1)}(x), f^{(n)}(x)) \le r^n d(f(x), x)/(1 - r)$ and since d and f are continuous, $d(f(x'), x') \le \limsup_n r^n d(f(x), x)/(1 - r) = 0$. Thus $f(x') = x'$.

For uniqueness, let $f(x'') = x''$ also. Then $d(x'', x') = d(f(x''), f(x'))$ since f fixes x' and x'', and $d(f(x''), f(x')) \le r d(x'', x')$ by the contraction property. Then $(1 - r)d(x'', x') \le 0$ and we conclude that $d(x'', x') = 0$. Thus $x'' = x'$.

9. If no point is isolated, each one-point set is closed nowhere dense. The countable union of these sets is the whole space, in contradiction to the Baire Category Theorem. An alternative argument is to appeal to Problem 2.

10. The set is closed and bounded, hence compact, and it is pathwise connected, hence connected. It is not, however, locally connected. Take, for example, the point $p = [c, 1/2]$ in X, where c is in C. The open ball of radius 1/4 around p has the property that no open subneighborhood of p is connected.

11. Fix x_0 in X, and let U be the set of all points in X that can be connected to x_0 by paths. The set U is nonempty, and we prove that it is open and closed. Being connected, it must then be all of X. It is open because the local pathwise connectedness means that any x in U can be connected to every point in some neighborhood of x by a path; hence U contains a neighborhood of each of its points and is open. To see that U is closed, let y be a limit point of U. If V is a pathwise connected open neighborhood of y, the set $U \cap V$ is nonempty because y is a limit point of U. Let z

be in $U \cap V$. Then x_0 can be connected to z by a path because of the defining property of U, and z can be connected to y by a path because V is pathwise connected. Hence x_0 can be connected to y by a path, and y is in U.

12. Any open subset of $\mathbb{R}^n$ is locally pathwise connected. So the desired conclusion follows from the previous problem.

13. Let the open set be U. For each x in U, let U_x be the union of all connected subsets of U containing x. It was shown in Section 8 that this is connected. For x and y in U, either $U_x = U_y$ or $U_x \cap U_y = \varnothing$ for the same reason. Then U is the disjoint union of its subsets U_x, which are connected. These are intervals, being connected, and they must be open in order not to be contained in larger connected subsets of U.

14. Same as for Proposition 2.21.

15. Suppose $\{f_t\}$ is totally bounded. Let $\epsilon > 0$ be given. Find, by total boundedness, real numbers $t_1, \dots, t_n$ such that for any t, there is an index $j = j(t)$ with $\|f_t - f_{t_j}\| < \epsilon$. Put $L/2 = \max\{|t_1|, \dots, |t_n|\}$. If we are given an interval of length $\geq L$, take t to be its center, so that the interval contains $[t - L/2, t + L/2]$. Choose j by total boundedness with $\|f_t - f_{t_j}\} < \epsilon$. Then $\|f_{t-t_j} - f_0\| < \epsilon$. So $t - t_j$ is an ϵ almost period, and this lies in $[t - L/2, t + L/2]$. Thus the Bohr condition holds.

Conversely suppose that the Bohr condition holds and f is uniformly continuous. Let $\epsilon > 0$ be given, and find L as in the Bohr condition for $\epsilon/2$ almost periods. Also, find some δ for uniform continuity of f and the number $\epsilon/2$. Choose $t_1, \dots, t_n$ in $I = [-L/2, L/2]$ such that any point in I is within δ of one of $t_1, \dots, t_n$. Let us see that the open balls of radius ϵ around $f_{t_1}, \dots, f_{t_n}$ together cover the set $\{f_t\}$ of all translates. If t is given, find an $L/2$ almost period $t - s$ in $[t - L/2, t + L/2]$. Here $|s| < L/2$, so that $\|f_{t-s} - f_0\| < \epsilon/2$ and $\|f_t - f_s\| < \epsilon/2$. Since $|s - t_j| < \delta$ for some j, we have $\|f_s - f_{t_j}\| < \epsilon/2$ by uniform continuity. Thus $\|f_t - f_{t_j}\| < \epsilon$.

16. Let T_f be the closure of the set of translates of f. This is complete by Problem 14. Theorem 2.36 shows that T_f is compact if and only if every sequence in it has a convergent subsequence, and this is the definition of Bochner almost periodicity. Theorem 1.46 shows that T_f is compact if and only if it is totally bounded, and this is equivalent to Bohr almost periodicity by Problem 15.

17. This is easier with the Bochner definition. For an example of closure under the various operations, consider closure under multiplication. Suppose that f and g are given and that we want a convergent subsequence from the sequence of translates $(fg)_{t_n}$. First choose a subsequence of $\{t_n\}$ such that those translates of f converge uniformly, and then choose a subsequence of that such that the translates of g converge uniformly. These sequences of translates of f and g will be uniformly bounded, and then it follows that the sequence of products converges uniformly.

For closure under uniform limits, we argue similarly with translates of each of the functions $\{f_n\}$ when $\lim f_n = f$ uniformly. A Cantor diagonal process is used to extract the sequence of translates to use for f.

18. If $\epsilon > 0$ is given, let U_n be the set where $|f(x) - f_n(x)| < \epsilon$. This is open by the assumed continuity, and $\bigcup_{n=1}^{\infty} U_n = X$ by the assumed convergence. Since

X is compact, some finite collection of U_n's suffices. Since the f_n's are pointwise increasing with n, the U_n's are increasing, and thus $X = U_N$ for some N. For that N, $|f(x) - f_N(x)| < \epsilon$. Then $|f(x) - f_n(x)| < \epsilon$ for $n \geq N$ since the f_n's are pointwise increasing.

19. If $0 \leq P_n(x) \leq \sqrt{x} \leq 1$, then $x \geq P_n(x)^2$ and the recursion shows that $P_{n+1}(x) \geq P_n(x)$. Also, $P_{n+1}(x) = P_n(x) + \frac{1}{2}(\sqrt{x} + P_n(x))(\sqrt{x} - P_n(x)) \leq P_n(x) + \frac{1}{2}(1+1)(\sqrt{x} - P_n(x)) = \sqrt{x}$.

20. By Problem 19, $P_n(x)$ increases pointwise to some $f(x)$. Passing to the limit in the recursion gives $f(x) = f(x) + \frac{1}{2}(x - f(x)^2)$, and thus $f(x)^2 = x$ and $f(x) = \sqrt{x}$. Since $\sqrt{x}$ is continuous and $[0,1]$ is compact, Dini's Theorem (Problem 18) shows that the convergence is uniform.

21. If x and y are given, then we are given three relevant functions in $\mathcal{A}$, possibly not all distinct. They are h_1 with $h_1(x) \neq h_1(y)$, h_2 with $h_2(x) \neq 0$, and h_3 with $h_3(y) \neq 0$. If $h_1(x)$ or $h_1(y)$ is 0, we can add a multiple of h_2 or h_3 to h_1 to obtain an h_4 with $h_4(x) \neq h_4(y)$, $h_4(x) \neq 0$, and $h_4(y) \neq 0$. The restrictions of h_4 and h_4^2 to the two-element set $\{x, y\}$ are linearly independent and therefore form a basis for the 2-dimensional space of restrictions. Hence some linear combination of h_4 and h_4^2 equals the given f at x and y.

22. Let f be in $C_{\mathbb{R}}(S)$ with $f(s_0) = 0$. Since $\mathcal{B}^{\mathrm{cl}} = C_{\mathbb{R}}(S)$, there exists a sequence $\{g_n\}$ in $\mathcal{B}$ with $\lim g_n = f$ uniformly. Then $\lim g_n(s_0) = f(s_0) = 0$ in particular. Put $f_n(s) = g_n(s) - g_n(s_0)$. Then $f_n(s_0) = 0$. The inequality $|f_n(s) - f(s)| = |g_n(s) - f(s) - g_n(s_0)| \leq |g_n(s) - f(s)| + |g_n(s_0)|$ shows that $\{f_n\}$ converges uniformly to f. The members of $\mathcal{A}$ are the members of $\mathcal{B}$ that vanish at s_0. The functions f_n have this property, and thus $\{f_n\}$ is a sequence in $\mathcal{A}$ converging uniformly to f.

24. For (a), we identify $C_0([0, +\infty), \mathbb{R})$ with the subalgebra of $C([0, +\infty], \mathbb{R})$ of continuous functions equal to 0 at $+\infty$. The function e^{-x} separates points on $[0, +\infty]$. Apply Problem 22 to the algebra it generates, namely the algebra of all finite linear combinations of e^{-nx} for n a positive integer.

For (b), let $\epsilon > 0$ be given, and choose $g(x) = \sum c_n e^{-nx}$ by (a) such that $\sup_{0 \leq x < +\infty} |f(x) - g(x)| \leq \epsilon$. The hypothesis forces $\int_0^b f(x) g(x)\, dx = 0$, and this is $\int_0^b f(x)^2\, dx - \int_0^b f(x)\big(f(x) - g(x)\big)\, dx$. Thus

$$0 \geq \int_0^b f(x)^2\, dx - \Big| \int_0^b f(x)\big(f(x) - g(x)\big)\, dx \Big|.$$

So $\int_0^b f(x)^2\, dx \leq \epsilon \int_0^b |f(x)|\, dx$. Since ϵ is arbitrary, $\int_0^b f(x)^2\, dx = 0$. Therefore $f = 0$.

25. Isometries are uniformly continuous. Applying Proposition 2.47 to the uniformly continuous function $\varphi_2 \circ \big(\varphi_1^{-1}\big|_{\varphi_1(X)}\big)$ of the dense subset $\varphi_1(X)$ of X_1^* into X_2^*, we obtain an isometry $\Psi : X_1^* \to X_2^*$ extending $\varphi_2 \circ \big(\varphi_1^{-1}\big|_{\varphi_1(X)}\big)$. Reversing the roles

of X_1^* and X_2^*, we obtain an isometry $\Phi : X_2^* \to X_1^*$ extending $\varphi_1 \circ \big(\varphi_2^{-1}\big|_{\varphi_2(X)}\big)$. Then $\Phi \circ \Psi$ is a continuous extension of the composition $\varphi_1 \circ \big(\varphi_2^{-1}\big|_{\varphi_2(X)}\big) \circ \varphi_2 \circ \big(\varphi_1^{-1}\big|_{\varphi_1(X)}\big)$, which is the identity map on $\varphi_1(X)$. Hence $\Phi \circ \Psi$ is the identity on X_1^*. Similarly $\Psi \circ \Phi$ is the identity on X_2^*. Thus Ψ is onto. This proves existence.

For uniqueness let Ψ and Ψ^* be two such maps. Then $\Psi^{-1} \circ \Psi^*$ is a continuous extension of the identity map on the dense subset $\varphi_1(X)$ of X_1^*, and hence it is the identity. Therefore $\Psi = \Psi^*$.

26. Theorem 2.60 says that X is dense in X^*. Then $X = X^*$ if and only if X is closed, and this happens if and only if X is complete, by Proposition 2.43.

27. The only one of these that requires explanation is (iv). We may assume that none of r, s, and $r + s$ is 0. Write $r = mp^k/n$ and $s = up^l/v$ with p not dividing any of r, s, u, v. Without loss of generality, we may assume $k \leq l$, so that $\max\{|r|_p, |s|_p\} = |r|_p = p^{-k}$. We have

$$r + s = mp^k/n + up^l/v = p^k\big(\tfrac{m}{n} + \tfrac{up^{l-k}}{v}\big) = p^k\big(\tfrac{mv+p^{l-k}nu}{nv}\big).$$

The denominator nv is not divisible by p. The part of the numerator within the parentheses is an integer, and we factor out any factors of p from it as p^a with $a \geq 0$. Then we have $|r + s|_p = p^{-(k+a)}$ and this is $\leq p^{-k}$ as required.

28. For the triangle inequality, let r, s, t be given. Then Problem 27 gives $d(r, t) = |r - t|_p = |(r - s) + (s - t)|_p \leq \max\{|r - s|_p, |s - t|_p\} \leq |r - s|_p + |s - t|_p = d(r, s) + d(s, t)$.

29. Part (a) will be illustrated by the more difficult (b) and (c). Multiplication by a member r of $\mathbb{Q}$ is a uniformly continuous function from $\mathbb{Q}$ into $\mathbb{Q}_p$; in fact, the equality $|r(s - s_0)|_p = |r|_p|s - s_0|_p$ shows that if ϵ is given, then the δ of uniform continuity can be taken as $|r|_p^{-1}\epsilon$. Proposition 2.47 then tells us how to form products rs for r in $\mathbb{Q}$ and s in $\mathbb{Q}_p$. For fixed s, the result is a uniformly continuous map of $\mathbb{Q}$ into $\mathbb{Q}_p$ since $|\cdot|_p$ extends continuously to $\mathbb{Q}_p$ and we have $|(r - r_0)s|_p = |r - r_0|_p|s|_p$. A second application of Proposition 2.47 extends the operation to a mapping of $\mathbb{Q}_p \times \mathbb{Q}_p$ into $\mathbb{Q}_p$ that is uniformly continuous in each variable when the other variable is held fixed. In fact, it is continuous in both variables since $|rs - r_0s_0|_p = |(r - r_0)s + r_0(s - s_0)|_p \leq |r - r_0|_p|s|_p + |r_0|_p|s - s_0| \leq |r - r_0|_p|s - s_0|_p + |r - r_0|_p|s_0|_p + |r_0|_p|s - s_0|$.

For (c), take a shell $A_{kn} = \big\{r \in \mathbb{Q}_p \bigm| p^{-k} \leq |r|_p \leq p^n\big\}$. This is a closed subset of $\mathbb{Q}_p$, hence complete. Reciprocal is a mapping from $A_{nk} \cap \mathbb{Q}$ into A_{kn} that is uniformly continuous because r and s in $A_{nk} \cap \mathbb{Q}$ implies $|r^{-1} - s^{-1}|_p = |(s - r)/rs|_p = |s - r|_p|r|_p^{-1}|s|_p^{-1} \leq p^{2n}|s - r|_p$. Hence reciprocal extends to a uniformly continuous mapping from A_{nk} to A_{kn}. These mappings are consistent as n and k tend to infinity, and thus reciprocal is a well-defined function from $\mathbb{Q}_p^\times$ to itself. It is continuous because the same computation as just given shows that $|r^{-1} - r_0^{-1}|_p = |r - r_0|_p|r|_p^{-1}|r_0|_p^{-1}$. If we write $|r|_p \geq \big||r_0|_p - |r - r_0|_p\big|$ and require that $|r - r_0|_p \leq \frac{1}{2}|r_0|_p$, then $|r^{-1} - r_0^{-1}|_p = |r - r_0|_p\big(\frac{1}{2}|r_0|_p\big)^{-1}|r_0|_p^{-1}$, and continuity of reciprocal at r_0 follows.

The abelian group axioms in (c) are associativity, commutativity, existence of the two-sided identity 1, and existence of two-sided reciprocals. To complete (c), we need associativity and commutativity. We can regard associativity as asserting the equality of two continuous functions from $\mathbb{Q}_p \times \mathbb{Q}_p \times \mathbb{Q}_p$ to $\mathbb{Q}_p$. These are equal on $\mathbb{Q} \times \mathbb{Q} \times \mathbb{Q}$, and this subset is dense. Hence the two functions are equal everywhere. Commutativity is proved similarly.

The distributive law in (d) is proved by the same technique used for associativity in (c). Thus $\mathbb{Q}_p$ is a field.

30. For (a), it is enough to prove that $S = \{t \in \mathbb{Q} \mid |t|_p \le 1\}$ is totally bounded. For x in $\mathbb{Q}$, let $C(\delta; x) = \{t \in \mathbb{Q} \mid |t - x|_p \le \delta\}$. It is enough to show for each integer $l \ge 0$ that $S \subseteq \bigcup_{r=0}^{p^l-1} C(p^{-l}; r)$. If t is given in S, t is of the form $t = m/n$ with m and n in $\mathbb{Z}$ and n nondivisible by p. Let n^{-1} denote the integer from 0 to $p^l - 1$ such that $nn^{-1} \equiv 1 \bmod p^l$, and let r denote the integer from 0 to $p^l - 1$ such that $n^{-1}m \equiv r \bmod p^l$. Then $m - nr \equiv 0 \bmod p^l$, and so $|m - nr|_p \le p^{-l}$. Since $|n|_p = 1$, $\left|\frac{m}{n} - r\right|_p \le p^{-l}$. Thus t is in $C(p^{-l}; r)$.

For (b), compact sets are closed and bounded by Proposition 2.34a. Conversely let E be closed and bounded. The set $T = \{t \in \mathbb{Q}_p \mid |t|_p \le 1\}$ is certainly closed. Since $\mathbb{Q}_p$ is complete, T is complete. Part (a) shows that T is totally bounded. By Theorem 2.46, T is compact. The given set E is contained in some set $T_n = \{t \in \mathbb{Q}_p \mid |t|_p \le p^n\}$. Multiplication by the member p^{-n} of $\mathbb{Q}_p$ carries T continuously onto T_n, and T_n is compact by Proposition 2.38. Since E is a closed subset of the compact set T_n, Proposition 2.34b shows that E is compact.

31. The first two assertions are routine consequences of (ii), (iii), and (iv). Let us consider the quotient $\mathbb{Z}_p/P$. We show that P is a maximal ideal. In fact, if I is an ideal in $\mathbb{Z}_p$ properly containing P, then I contains some element t with $|t|_p = 1$. Then (iii) shows that t^{-1} has $|t^{-1}|_p = 1$ and lies in $\mathbb{Z}_p$. Since t is in I and t^{-1} is in $\mathbb{Z}_p$, their product 1 is in I. Thus $I = \mathbb{Z}_p$. In other words, P is a maximal ideal. Hence $\mathbb{Z}_p/P$ is a field. To complete the argument, we show that $\mathbb{Z}_p/P$ has exactly p elements. Given x in $\mathbb{Z}_p$, choose m/n in $\mathbb{Q}$ with $\left|x - \frac{m}{n}\right|_p \le p^{-1}$, by denseness of $\mathbb{Q}$ in $\mathbb{Q}_p$. Here $\left|\frac{m}{n}\right|_p \le 1$, and we may assume that n is nondivisible by p. Arguing as in Problem 30a, we can find r in $\{0, 1, \dots, p-1\}$ such that $\left|\frac{m}{n} - r\right|_p \le p^{-1}$. Then $|x - r|_p \le \max\left\{\left|x - \frac{m}{n}\right|_p, \left|\frac{m}{n} - r\right|_p\right\} \le p^{-1}$ by the ultrametric inequality. So $x = (x-r)+r$ with $x-r$ in P. Thus $\{0, 1, \dots, p-1\}$ represents all cosets of $\mathbb{Z}_p/P$. Finally no two distinct elements r and r' in $\{0, 1, \dots, p-1\}$ have $|r - r'|_p \le p^{-1}$ because this inequality would entail having $r - r'$ divisible by p.

Chapter III

1. For (a), $|TS|^2 = \sum_j |TS(e_j)|^2 = \sum_j |\sum_i (S(e_j), e_i)T(e_i)|^2$. Use of the triangle inequality and then the Schwarz inequality shows that this expression is $\le$

$\sum_j \big(\sum_i |(S(e_j), e_i)|\,|T(e_i)|\big)^2 \le \sum_j \Big(\big(\sum_i |(S(e_j), e_i)|^2\big)^{1/2}\big(\sum_i |T(e_i)|^2\big)^{1/2}\Big)^2 = \sum_j |S(e_j)|^2|T|^2 = |S|^2\,|T|^2$. Part (b) is routine.

2. The member of $L(\mathbb{R}^n, \mathbb{R}^m)$ with matrix A.

3. $\limsup_{h\to 0} \big(|h|^{-1}|f(h) - 0 - 0|\big) \le \limsup_{h\to 0} (|h|^{-1}|h|^2) = 0$.

4. The formula is $\frac{d}{dt} f(x + tu)\big|_{t=0} = \sum_j u_j \frac{\partial f}{\partial x_k}(x)$. The argument is written out within the proof of Theorem 3.11.

5. $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$, $e^t\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$, $\begin{pmatrix} \cos t & i\sin t \\ i\sin t & \cos t \end{pmatrix}$, $\begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$.

7. The equality is false because the left side is positive and the right side is negative. In fact, the left side is $\int_0^1 \big[\lim \int_1^N (e^{-xy} - 2e^{-xy})\,dx\big]\,dy$, which equals $\int_0^1 \lim \big[-e^{-xy}/y + e^{-2xy}/y\big]_1^N\,dy = \int_0^1 \frac{1}{y}\big[e^{-y} - e^{-2y}\big]\,dy$; since $e^{-y} > e^{-2y}$ on $(0, 1)$, the left side is > 0. Meanwhile, the right side is $\int_1^\infty \big[-e^{-xy}/x + e^{-2xy}/x\big]_0^1\,dx = \int_1^\infty \frac{1}{x}\big[e^{-2x} - e^{-x}\big]\,dx$; since $e^{-2x} < e^{-x}$ on $(1, \infty)$, the right side is < 0.

8. Define $\|\cdot\|_2$ as in Section I.10, and let $f_x(t) = f(x - t)$; the latter definition is not the one used earlier in the book. For (a), the Schwarz inequality gives

$$\begin{aligned} |f * g(x) - f * g(x_0)| &= \big|\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi}[f(x - t) - f(x_0 - t)]g(t)\,dt\big| \\ &= \|f_x - f_{x_0}\|_2\|g\|_2 \le \|g\|_2 \sup_t |f(x - t) - f(x_0 - t)|, \end{aligned}$$

and the right side tends to 0 as x tends to x_0 by uniform continuity of f. This proves that $f * g$ is continuous. The periodicity is evident. The proof that $f * g = g * f$ is the same as the proof in Section I.10 that $f * D_N = D_N * f$.

For (b), an application of Fubini's Theorem (Corollary 3.33) and a change of variables gives $\frac{1}{2\pi}\int_\pi^\pi f * g(x)e^{-inx}\,dx = \big(\frac{1}{2\pi}\big)^2 \int_\pi^\pi \int_\pi^\pi f(x - t)g(t)e^{-inx}\,dt\,dx = \big(\frac{1}{2\pi}\big)^2 \int_\pi^\pi \int_\pi^\pi f(x - t)g(t)e^{-inx}\,dx\,dt = \big(\frac{1}{2\pi}\big)^2 \int_\pi^\pi \int_\pi^\pi f(x)g(t)e^{-in(x+t)}\,dx\,dt = \big(\frac{1}{2\pi}\big)^2 \int_\pi^\pi \int_\pi^\pi f(x)g(t)e^{-inx}e^{-int}\,dx\,dt = c_n d_n$.

For (c), we apply the Weierstrass M test. It is enough to prove that $\sum_n |c_n d_n| < +\infty$, and the Schwarz and Bessel inequalities together do this:

$$\sum_n |c_n d_n| \le \Big(\sum_n |c_n|^2\Big)^{1/2}\Big(\sum_n |d_n|^2\Big)^{1/2} \le \|f\|_2\|g\|_2 < +\infty.$$

9. Write out each side as an iterated integral, and apply Fubini's Theorem (Corollary 3.33).

10. For the partial derivatives, $\frac{\partial x}{\partial x}(0, 0) = \frac{d}{dx} f\big(\frac{x0}{x^2+0}\big)\big|_{x=0} = 0$ and $\frac{\partial f}{\partial y}(0, 0) = 0$ similarly. The fact that f is not continuous at $(0, 0)$ is a special case of Problem 11a.

11. For (a), the homogeneity says in particular that $f(rx) = f(x)$ for $r > 0$ and $|x| = 1$. Then $\sup_{y\ne 0} |f(y)| = \sup_{|x|=1} |f(x)|$, and the right side is finite, being the maximum value of a continuous function on a compact set. If $f(y)$ is continuous at

$y = 0$, then $f(0) = \lim_{r\downarrow 0} f(rx) = f(x)$ for every x with $|x| = 1$ and so f must be constantly equal to $f(0)$.

For (b), $\limsup_{rx\to 0} |f(rx)| = \limsup_{rx\to 0} r^d |f(x)| = 0$ if $d > 0$ since $f(x)$ is bounded for $|x| = 1$. Thus f is continuous at 0 if $d > 0$ and $f(0) = 0$. If $d < 0$, then $\limsup_{rx\to 0} r^d|f(x)| = +\infty$ if $d < 0$ and $f(x) \neq 0$.

For (c), we have $f(rx) = r^d f(x)$ for any $x = (x_1, \dots, x_n) \neq 0$. Put $g = f \circ m_r$, where m_r refers to multiplication by r. The homogeneity gives $g = r^d f$, and thus $\frac{\partial g}{\partial x_j}(x) = r^d \frac{\partial f}{\partial x_j}(x)$. On the other hand, the chain rule gives $\frac{\partial g}{\partial x_j}(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(rx) \frac{\partial (rx_i)}{\partial x_j}(x) = r \frac{\partial f}{\partial x_j}(rx)$. So $r^d \frac{\partial f}{\partial x_j}(x) = r \frac{\partial f}{\partial x_j}(rx)$, and (c) follows.

For (d), the given conditions say that $f(tx) = tf(x)$ for all real t. Then $\frac{\partial f}{\partial x_j}(0) = \lim_{t\to 0} t^{-1}\big(f(0 + te_j) - 0\big) = \lim_{t\to 0} t^{-1} t f(e_j) = f(e_j)$. On the other hand, (c) says that $\partial f/\partial x_j$ is homogeneous of degree 0, and (a) says that $\partial f/\partial x_j$ cannot be continuous at 0 unless it is constant.

12. Part (a) follows from Problem 11b. In (b), $\frac{\partial f}{\partial x}(0) = \frac{d}{dt} f(0 + t(1,0))\big|_{t=0} = \frac{d}{dt} t\big|_{t=0} = 1$ and $\frac{\partial f}{\partial y}(0) = \frac{d}{dt} f(0 + ty)\big|_{t=0} = \frac{d}{dt} 0\big|_{t=0} = 0$. The failure of continuity is by parts (a) and (c) of Problem 11.

For (c), we have $\frac{d}{dt} f(0 + tu)\big|_{t=0} = \frac{d}{dt} t \cos^3\theta\big|_{t=0} = \cos^3\theta$. If f were differentiable at $x = 0$, the chain rule would give $\frac{d}{dt} f(0 + tu)\big|_{t=0} = u_1 \frac{\partial f}{\partial x}(0) + u_2 \frac{\partial f}{\partial y}(0) = \cos\theta$. Since $\cos^3\theta$ is not identically equal to $\cos\theta$, f is not differentiable at 0.

13. Part (a) follows from (a), (b), and (c) of Problem 11. About 0, the function f is even in x and even in y, and hence the first partial derivatives are odd about 0. Then part (b) follows from Problem 11d. To calculate the results for (c), we need to compute $\frac{\partial f}{\partial y}(x, 0)$ for $x \neq 0$ and $\frac{\partial f}{\partial x}(0, y)$ for $y \neq 0$. The first of these is x, and the second is $-y$. The formulas for the second partial derivatives follow.

14. For $n \geq 0$, $r^n e^{in\theta} = (x+iy)^n$ is of class C^∞, and so is $r^n e^{-in\theta} = (x-iy)^n$. For the first of these functions, $\frac{\partial^2}{\partial x^2}(x+iy)^n = n(n-1)(x+iy)^{n-2}$, while $\frac{\partial^2}{\partial y^2}(x+iy)^n = i^2 n(n-1)(x+iy)^{n-2}$. Hence $\Delta(x+iy)^n = 0$. The result for $(x-iy)^n$ follows by taking complex conjugates. The final conclusion is a routine consequence of Theorem 1.37, the complex-valued version of Theorem 1.23, and the fact that each term is harmonic.

15. This follows by direct calculation.

16. In the notation of Theorem 3.17, $\varphi(x, y)$ is $\begin{pmatrix} x^4 y + x \\ x + y^3 \end{pmatrix}$, a is $(1, 1)$, and b is $(2, 2)$. One checks that $\varphi'(1, 1) = \begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}$. The locally defined inverse function f near $(2, 2)$ has $f'(2, 2) = \varphi'(1, 1)^{-1} = \begin{pmatrix} 3/14 & -1/14 \\ -1/14 & 5/14 \end{pmatrix}$, and $\frac{\partial F}{\partial u}(2, 2)$ is the upper left entry of this, namely $3/14$.

17. All 6 derivatives of possible interest are given by the matrix product

$\begin{pmatrix} 2 & -1 & 0 \\ 2 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 0 & -\pi/2 \\ 0 & 0 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 0 & -\pi/2 \\ 0 & -\pi \\ 0 & 3\pi/2 \end{pmatrix}$. Then $\frac{\partial x}{\partial u}(\pi/2, 0) = 0$ and $\frac{\partial x}{\partial v}(\pi/2, 0) = -\pi/12$. The function $x(u, v)$ is of class C^∞ by Corollary 3.21.

18. The map in question is $X \mapsto X^2$ and is the composition of $X \mapsto (X, X)$ followed by $(U, V) \mapsto UV$. Here we can write $UV = L(U)V = R(V)U$, where $L(U)$ is the linear function "left multiplication by U" on matrix space and $R(V)$ is the linear function "right multiplication by V." The derivative of $(U, V) \mapsto UV$ is then $(\, R(V) \quad L(U)\,)$ by Problem 2. Hence the derivative of $X \mapsto X^2$, by the chain rule, is

$$(\, R(V) \quad L(U)\,)\begin{pmatrix} 1 \\ 1 \end{pmatrix}\Bigg|_{U=V=X} = (R(V) + L(U))\big|_{U=V=X} = R(X) + L(X).$$

At $X = 1$, this is $R(1) + L(1)$, which is "multiplication by 2" and is invertible. The Inverse Function Theorem thus applies.

19. We may assume that $g'(x_0) \neq 0$, thus that $\frac{\partial g}{\partial x_i}(x_0) \neq 0$ for some i. We take this i to be $i = n$; the other cases involve only notational changes. Write $x = (x', x_n)$ with $x' \in \mathbb{R}^{n-1}$, and write $x_0 = (a, b)$ similarly. Then the Implicit Function Theorem produces a real-valued C^1 function $h(x')$ defined on an open set V about the point a in $\mathbb{R}^{n-1}$ such that $h(a) = b$, $g(x', h(x')) = 0$ for all x' in V, and $\frac{\partial h}{\partial x_j}(a) = -\big(\frac{\partial g}{\partial x_n}(a, b)\big)^{-1}\big(\frac{\partial g}{\partial x_j}(a, b)\big)$ for $1 \leq j < n$. Let $H(x) = (x', h(x'))$. Form $f \circ H$, which has a local maximum or minimum at $x' = a$ in V. All the first partial derivatives of this function must be 0 at $x' = a$. Thus, for $1 \leq j \leq n - 1$, $0 = \frac{\partial (f \circ H)}{\partial x_j}(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0)\frac{\partial H_i}{\partial x_j}(a)$. Since $H_i(x) = x_i$ for $i < n$, all the terms of this sum are 0 except possibly for the j^{th} and the n^{th}. Thus $0 = \frac{\partial f}{\partial x_j}(x_0) + \frac{\partial f}{\partial x_n}(x_0)\frac{\partial h}{\partial x_j}(a)$ $= \frac{\partial f}{\partial x_j}(x_0) - \big(\frac{\partial f}{\partial x_n}\big)(x_0)\big(\frac{\partial g}{\partial x_n}(a, b)\big)^{-1}\big(\frac{\partial g}{\partial x_j}(a, b)\big)$ for $j < n$. The right side is 0 trivially for $j = n$, and thus the result follows with $\lambda = -\big(\frac{\partial g}{\partial x_n}(a, b)\big)^{-1}$.

20. The difficulty in handling this inequality as a maximum-minimum problem is the question of existence. Lagrange multipliers can constrain matters to a compact set, and then existence is no longer an obstacle. The domain D initially will be the set where $a_1 \geq 0, \dots, a_n \geq 0$. Fix a number c, and let $g(a_1, \dots, a_n) = \frac{1}{n}(a_1 + \cdots + a_n) - c$ and $f(a_1, \dots, a_n) = \sqrt[n]{a_1 \cdots a_n}$. The subset of D where $g(a_1, \dots, a_n) = 0$ is compact, and f must have an absolute maximum on it. This maximum cannot occur where any a_j equals 0 since f is 0 at such points. So it is at a point in the set U where all a_j are > 0. Apply Lagrange multipliers on U. The resulting equations are $\frac{1}{n}(a_1 \cdots a_n)^{1/n}\big/a_j = 1/n$ for $1 \leq j \leq n$, as well as the constraint equation $\frac{1}{n}(a_1 + \cdots + a_n) = c$. The first n equations show that all a_j's must be equal, and the constraint equation shows that they must equal c. The desired inequality is true in this case and hence is true in all cases.

Chapter IV

1. For (a), $\frac{1}{2}y^2 = -\frac{1}{2}t^2 + c$. Adjusting c, we have $y^2 = -t^2 + c$. Then $y(t) = \pm\sqrt{c - t^2}$. For (b), the exceptional points are $(t_0, 0)$. For (c), a solution with $y(t_0) = y_0$ is $y(t) = \operatorname{sgn}(y_0)\sqrt{y_0^2 + t_0^2 - t^2}$.

2. In Theorem 4.1, take $a = 1$ and $b = 1$. Then $M = 2$ and $a' = \frac{1}{2}$. The theorem therefore gives a solution for $|t| < 1/2$.

3. To be an integral curve, $(x(t), y(t))$ must satisfy $x'(t) = \sqrt{x}$ and $y'(t) = 1/2$. Then $2\sqrt{x(t)} = t + c_1$ and $y(t) = \frac{1}{2}t + c_2$. At some unspecified time t_0, the curve is to pass through $(1, 1)$. Then $x(t_0) = 1$ and $y(t_0) = 1$; these force $2 = t_0 + c_1$ and $1 = \frac{1}{2}t_0 + c_2$. So $(x(t), y(t)) = \big(\frac{1}{4}(t - t_0 + 2)^2, \frac{1}{2}(t - t_0 + 2)\big)$. If $t_0 = 0$, for example, the curve is $(x(t), y(t)) = \big(\frac{1}{4}(t + 2)^2, \frac{1}{2}(t + 2)\big)$.

4. This uses the multivariable chain rule, Proposition 3.28b, and the Fundamental Theorem of Calculus. The derivative in question is

$$= (2t)(1/t^2)\sin(t^3) + \int_0^{t^2} (\partial/\partial t)(s^{-1}\sin(st))\,ds = (2/t)\sin(t^3) + \int_0^{t^2} \cos(st)\,ds$$
$$= (2/t)\sin(t^3) + \big[t^{-1}\sin(st)\big]_{s=0}^{t^2} = (2/t)\sin(t^3) + t^{-1}\sin(t^3).$$

5. $y(t) = 2 + c_1e^t + c_2e^{2t}$.

6. For (a), $J = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$ for the first, and $J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}$ and $B = \begin{pmatrix} 0 & i & -i \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ for the second. For (b), the bases are $e^{3t}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $e^{3t}\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} + t\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right)$ for the first, and $e^t\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e^{it}\begin{pmatrix} i \\ 0 \\ 1 \end{pmatrix}, e^{-it}\begin{pmatrix} -i \\ 0 \\ 1 \end{pmatrix}$ for the second.

Part (b) can be solved directly without solving part (a) first. Consider the 2-by-2 example. The only root of the characteristic polynomial is 3, and it has multiplicity 2. We solve $(A - 3\cdot 1)k_0 = 0$ and get $k_0 = \begin{pmatrix} c \\ 2c \end{pmatrix}$. Then we solve $(A - 3\cdot 1)l_0 = \begin{pmatrix} c \\ 2c \end{pmatrix}$ and get $l_0 = \begin{pmatrix} d \\ c + 2d \end{pmatrix}$. Choose any $c \neq 0$ and any d, say $c = 1$ and $d = 0$. Then $k_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, and $l_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and we obtain the solutions in the form given above. For more complicated examples, the choice of these constants can get tricky, but this method works quickly for easy examples.

7. For $n = 1$, $\det(\lambda - (-a_0)) = \lambda + a_0$. Assume the result for $n - 1$, and expand

the n^{th}-order determinant by cofactors about the first column. Then

$$\begin{aligned}
\det(\lambda 1 - A) &= \det\begin{pmatrix} \lambda & -1 & 0 & 0 & \cdots & 0 & 0 \\ & \lambda & -1 & 0 & \cdots & 0 & 0 \\ & & \lambda & -1 & \cdots & 0 & 0 \\ & & & \ddots & \ddots & \vdots & \vdots \\ & & & & \lambda & -1 & 0 \\ & & & & & \lambda & -1 \\ a_0 & a_1 & a_2 & & \cdots & & a_{n-1} \end{pmatrix} \\
&= \lambda \det\begin{pmatrix} \lambda & -1 & 0 & \cdots & 0 & 0 \\ & \lambda & -1 & \cdots & 0 & 0 \\ & & \ddots & \ddots & \vdots & \vdots \\ & & & \lambda & -1 & 0 \\ & & & & \lambda & -1 \\ a_1 & a_2 & & \cdots & & a_{n-1} \end{pmatrix} + (-1)^{n-1} a_0 \det\begin{pmatrix} -1 & & & \\ & -1 & \cdots & 0 \\ & & \ddots & \\ & * & \cdots & \ddots \\ & & & -1 \end{pmatrix} \\
&= \lambda(\lambda^{n-1} + a_{n-1}\lambda^{n-2} + \cdots + a_1) + (-1)^{n-1} a_0 (-1)^{n-1} \\
&= \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0,
\end{aligned}$$

the next-to-last equality following by induction.

8. In (a), let $|f_n(t) \le M$ for all t and n. Then $|F_n(t) - F_n(t')| = \left|\int_{t'}^{t} f_n(s)\,ds\right| \le M|t - t'|$. Thus equicontinuity holds with $\delta = \epsilon/M$.

In (b), we solve the equation explicitly, using variation of parameters. The solutions of the homogeneous equation are $c_1 \cos t + c_2 \sin t$, and computation shows that the unique solution of the inhomogeneous equation with the given initial condition is $y^*(t) = -(\cos t)\int_0^t (\sin s) f(s)\,ds + (\sin t)\int_0^t (\cos s) f(s)\,ds$. Each integral is equicontinuous by the same argument as in (a), and the operations of multiplication by a bounded continuous function and addition preserve the equicontinuity.

In (c), we do not know explicit formulas for the solutions of the homogeneous equation, but the same argument as in (b) with variation of parameters will work anyway.

10. For any C^2 periodic function f, the n^{th} Fourier coefficient c_n of f has $|c_n| \le n^{-2} \sup|f''|$. The function $v(r, \theta)$, being a composition of two C^2 functions, is C^2 for $0 \le r < 1$ and $|\theta| \le \pi$, and hence $\sup\left|\frac{\partial^2 v}{\partial \theta^2}\right|$ is bounded by some M for $0 \le r \le 1-\delta$. Then we obtain $|c_n(r)| \le M/n^2$.

11. The function $(u \circ R_\varphi)(x, y)e^{-ik\varphi}$ is of class C^2 jointly in x, y, φ. By Proposition 3.28 we can pass the second derivatives with respect to x and y under the given integral sign with respect to φ. The integrand is harmonic in (x, y) for each φ, and therefore the integral itself is harmonic. The integral itself is given by

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} v(r, \theta + \varphi)e^{-ik\varphi}\,d\varphi = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_{n=-\infty}^{\infty} c_n(r)e^{in\theta}e^{i(n-k)\varphi}\,d\varphi.$$

The series in the integrand is uniformly convergent as a function of φ, by the estimate in Problem 10 and by the Weierstrass M-test. Theorem 1.31 says that we can interchange sum and integral, and then the right side above collapses to $c_k(r)e^{in\theta}$.

12. Starting from $v(r,\theta)=u(r\cos\theta, r\sin\theta)$, we compute $\frac{\partial v}{\partial r}$ and $\frac{\partial v}{\partial\theta}$ by the chain rule and obtain

$$\frac{\partial v}{\partial r}=\cos\theta\,\frac{\partial u}{\partial x}+\sin\theta\,\frac{\partial u}{\partial y}\quad\text{and}\quad\frac{\partial v}{\partial\theta}=-r\sin\theta\,\frac{\partial u}{\partial x}+r\cos\theta\,\frac{\partial u}{\partial y}.$$

Using the same technique, we form $\frac{\partial^2 v}{\partial r^2}$ and $\frac{\partial^2 v}{\partial\theta^2}$ in terms of the partial derivatives of u, and we find that

$$\Delta u=\frac{\partial^2 v}{\partial r^2}+\frac{1}{r}\frac{\partial v}{\partial r}+\frac{1}{r^2}\frac{\partial^2 v}{\partial\theta^2}.$$

Substituting $v(r,\theta)=c_k(r)e^{ik\theta}$ and taking into account that $\Delta u=0$, we obtain

$$0=e^{ik\theta}\big(c_k''+r^{-1}c_k'-k^2r^{-2}c_k\big).$$

Thus $r^2c_k''+rc_k'-k^2c_k=0$. This is an Euler equation. The solutions are $c_k(r)=a_kr^{|k|}+b_kr^{-|k|}$ if $k\neq 0$ and are $a_0+b_0\log r$ if $k=0$. Taking into account that $c_k(r)$ is differentiable at $r=0$, we obtain $c_k(r)=a_kr^{|k|}$ for all k. Substitution gives $v(r,\theta)=\sum_{n=-\infty}^{\infty}c_nr^{|n|}e^{in\theta}$.

13. Since $f_R(\theta)=\sum_{n=-\infty}^{\infty}c_nR^{|n|}e^{in\theta}$ and $P_{r/R}(\theta)=\sum_{n=-\infty}^{\infty}(r/R)^{|n|}e^{in\theta}$, the result follows immediately from Problem 8b at the end of Chapter III.

15. For (a), substitute $y=uv$, $y'=u'v+uv'$, and $y''=u''v+2u'v'+uv''$ into the equation for y, take into account that $u''+Pu'+Qu=0$, and get $2u'v'+uv''+Puv'=0$. Put $w=v'$. We can rewrite our equation as $w'=(-P-2u'/u)w$ since u is assumed nonvanishing. Then Problem 14 gives $w(t)=ce^{-\int P\,dt-2\int(u'/u)\,dt}=ce^{-\int P\,dt}e^{\log(|u|^{-2})}=cu(t)^{-2}e^{-\int P(t)\,dt}$.

For (b), the formula in (a) gives $v'(t)=ce^{-t^2/2}$, and hence $y(t)=u(t)v(t)=e^{t^2/2}\int_0^t e^{-s^2/2}\,ds$.

16. The substitution leads to $uv''+(2u'+Pu)v'+(u''+Pu'+Qu)v=0$. Thus the condition is $2u'+Pu=0$. By Problem 14, $u(t)$ is a multiple of $e^{-\int(P/2)\,dt}$. The computation of $R(t)$ is then routine.

17. Substitution of $v=ur^{-1/2}$ shows that $L(v)=r^{1/2}L_0(u)$ with L_0 of the indicated form.

18. For (a), the formula is $d_n=-\sum_{k=1}^n c_kd_{n-k}$, with $d_0=1$. For (b), we have $d_1=-c_1d_0=-c_1$, so that $|d_1|=|c_1|\le Mr^1$. Thus $|d_n|\le M(M+1)^{n-1}r^n$ for $n=1$. Assume that $|d_k|\le Mr^k$ for $1\le k<n$. Then $|d_n|\le\sum_{k=0}^{n-1}|c_{n-k}||d_k|\le|c_n|+\sum_{k=1}^{n-1}(Mr^{n-k})(M(M+1)^{k-1}r^k)\le Mr^n+M^2r^n\sum_{k=1}^{n-1}(M+1)^{k-1}$. This is

$$\begin{aligned}&=Mr^n\big(1+M\textstyle\sum_{k=1}^{n-1}(M+1)^{k-1}\big)\\&=Mr^n(1+M((M+1)^{n-1}-1)/((M+1)-1)\\&=Mr^n(1+(M+1)^{n-1}-1)=M(M+1)^{n-1}r^n.\end{aligned}$$

For (c), we may assume that $f(0) = 1$. Write $f(x) = \sum_{n=0}^{\infty} c_n x^n$, and define d_n as in the answer to (a). The estimate in (b) shows that the power series $g(x) = \sum_{n=0}^{\infty} d_n x^n$ has positive radius of convergence, and Theorem 1.40 shows that $f(x)g(x) = 1$ on the common region of convergence. Then $g(x) = 1/f(x)$, and $1/f(x)$ is exhibited as the sum of a convergent power series.

19. The indicial equation is $s(s-1) + a_0 s + b_0 = 0$, where $a_0 = P(0)$ and $b_0 = Q(0)$. Thus $s_1 + s_2 = 1 - a_0$.

In (a), we apply Problem 15a with $u(t) = t^{s_1} \sum_{n=0}^{\infty} c_n t^n$. The expression $P(t)$ in that problem has become $t^{-1}P(t)$ here, and we obtain $v'(t) = u(t)^{-2} e^{-\int t^{-1} P(t)\,dt}$. In the integrand of the exponent, we separate the term $-a_0/t$ from the power series, and we see that $v'(t) = u(t)^{-2} e^{-a_0 \log t} \times$ power series $= t^{-a_0} u(t)^{-2} \times$ power series, the power series having nonzero constant term since exponentials are nowhere vanishing. This is of the form $t^{-2s_1 - a_0} \times$ power series as a consequence of Problem 18 and Theorem 1.40, the power series having nonzero constant term. When this expression is integrated to form $v(t)$, the t^{-1} produces a logarithm, and the rest produces powers of t. Thus $v(t)$ equals $c \log t + t^{-2s_1 - a_0 + 1} \times$ power series; here the power series has nonzero constant term. Then $u(t)v(t) = cu(t) \log t + t^{s_1} t^{-2s_1 - a_0 + 1} \times$ power series; once again the power series has nonzero constant term. The exponent of t in the second term is $-s_1 + 1 - a_0 = -s_1 + (s_1 + s_2) = s_2$, and (a) is done.

In (b), we know that there is only one solution beginning with t^{s_1}, and thus we must have $c \neq 0$ in (a). Another way to see this conclusion is to recognize that the exponent of $t^{-2s_1 - a_0}$ in $v'(t)$ is just -1 since $2s_1 = s_1 + s_2$. Thus the coefficient of t^{-1} in integrating to form $v(t)$ is not 0, and the logarithm occurs.

In (c), we know from a computation in the text that no series solution begins with t^{-p} except when $p = 0$, and thus the first argument for (b) applies.

20. When $t = t_{k-1}$ is substituted into the formula valid for $t_{k-1} < t \leq t_k$, we get $y(t) = y(t_{k-1})$; so the formula is valid also at t_{k-1}.

We induct on k. For $k = 0$, $y(t_0) = y_0$. Assume inductively for $k > 0$ that $|y(t_{k-1}) - y(t_0)| \leq M|t_{k-1} - t_0| \leq Ma' \leq b$. For $t_{k-1} \leq t \leq t_k$, the displayed formula in the problem implies $|y(t) - y(t_{k-1})| = |F(t_{k-1}, y(t_{k-1}))|\,|t - t_{k-1}|$. Since $(t_{k-1}, y(t_{k-1}))$ lies in R', $|F|$ is $\leq M$ on it. Thus $|y(t) - y(t_{k-1})| \leq M|t - t_{k-1}| \leq Ma' \leq b$. If $t_{l-1} \leq t \leq t_l$, then adding such inequalities gives $|y(t) - y(t_0)| \leq M|t_1 - t_0| + \cdots + M|t_{l-1} - t_{l-2}| + M|t - t_{l-1}| = M|t - t_0|$ as required. Since $|t - t_0| \leq a'$, we have $M|t - t_0| \leq Ma' \leq b$. Thus $(t, y(t))$ is in R'.

21. We may assume that $t' \leq t$. If t' and t lie in the same interval $[t_{k-1}, t_k]$ of the partition, then $y(t) - y(t') = F(t_{k-1}, y(t_{k-1}))(t - t')$. Taking absolute values gives $|y(t) - y(t')| \leq M|t - t'|$.

Otherwise let $t' \leq t_l \leq t_{k-1} \leq t$. Then each pair of points (t', t_l),(t_l, t_{l+1}), ... ,(t_{k-2}, t_{k-1}),(t_{k-1}, t) lies in a single interval of the partition. Adding the estimates for each and taking into account that each difference of t values is ≥ 0, we obtain $|y(t) - y(t')| \leq M|t - t'|$.

22. Let $t_{k-1} \le t \le t_k$. Then $\int_{t_0}^{t} y'(s)\,ds = \sum_{j=1}^{k-1} \int_{t_{j-1}}^{t_j} y'(s)\,ds + \int_{t_{k-1}}^{t} y'(s)\,ds = (y(t_1) - y(t_0)) + \cdots + (y(t_{k-1}) - y(t_{k-2})) + (y(t) - y(t_{k-1})) = y(t) - y(t_0)$, by an application of the Fundamental Theorem of Calculus on each interval. If $t_{k-1} < s < t_k$, we have $|y'(s) - F(s, y(s))| = |F(t_{k-1}, y(t_{k-1})) - F(s, y(s))|$. Here $|s - t_{k-1}| \le |t_k - t_{k-1}| \le \delta$ by the choice of the partition. Again by the choice of the partition, $|y(s) - y(t_{k-1})| \le M|s - t_{k-1}| \le M(\delta/M) = \delta$. By the definition of δ in terms of ϵ and the uniform continuity of F, we conclude that $|y'(s) - F(s, y(s))| \le \epsilon$.

23. We have $\big|y(t) - \big(y_0 + \int_{t_0}^{t} F(s, y(s))\,ds\big)\big| = \big|\int_{t_0}^{t} [y'(s) - F(s, y(s))]\,ds\big| \le \int_{t_0}^{t} |y'(s) - F(s, y(s))|\,ds \le \int_{t_0}^{t} \epsilon\,ds \le \epsilon|t - t_0| \le \epsilon a'$.

24. The statement of Problem 21 proves uniform equicontinuity with $\delta = \epsilon/M$. If we specialize to $t' = t_0$, it implies uniform boundedness.

25. Let $y(t) = \lim y_{n_k}(t)$ uniformly. The functions $y_{n_k}(t)$ are continuous, and the uniform limit of continuous functions is continuous. Hence $y(t)$ is continuous. By Problem 23 we have $\big|y_{n_k}(t) - \big(y_0 + \int_{t_0}^{t} F(s, y_{n_k}(s))\,ds\big)\big| \le \epsilon_{n_k} a'$ for each k. We take the limsup of this expression as k tends to infinity. We know that $y_{n_k}(t)$ tends uniformly to $y(t)$. Then $y_{n_k}(s)$ tends uniformly to $y(s)$ uniformly for $t_0 \le s \le t$. By uniform continuity of F, $F(s, y_{n_k}(s))$ tends uniformly to $F(s, y(s))$. By Theorem 1.31, $\int_{t_0}^{t} F(s, y_{n_k}(s))\,ds$ tends to $\int_{t_0}^{t} F(s, y(s))\,ds$.

Chapter V

1. For (a) and (c), the answer is 2^k for $0 \le k \le n$. However, the assertion in (d) is false; for a counterexample, take $X = \{1, 2, 3, 4\}$, and let $\mathcal{B}$ consist of all sets with an even number of elements. For (b), the associativity is proved by observing that $A \,\Delta\, B \,\Delta C$ is the set of all elements that lie in an odd number of the sets A, B, C.

2. Let $X = \{1, 2, 3\}$ with the σ-algebra consisting of all subsets. Take $\rho(\{1\}) = \rho(\{3\}) = +2$, $\rho(\{2\}) = -3$, $A = \{1, 2\}$, and $B = \{2, 3\}$.

4. This can be worked out carefully, but it is easier to use Problem 3 and apply dominated convergence to see that the measure of the left side is $\limsup \mu(E_n)$, and the measure of the right side is $\liminf \mu(E_n)$.

5. Part (a) is proved the same way as for Lebesgue measure. In (b), the interval I of rationals from 0 to 1 has $\mu(I) = 1$, and it is a countable union of one-point sets $\{p\}$, each of which has $\mu(\{p\}) = 0$.

6. Argue by contradiction. If E^c is not dense, then there is a nonempty open interval U in [0, 1] with $U \cap E^c = \varnothing$ and hence $U \subseteq E$. Since $\mu(U) > 0$, we must have $\mu(E) > 0$.

7. As soon as $\sup \mu(A)$ is known to be finite, B can be constructed as the union of a sequence of sets whose measures increase to the supremum. Thus assume that the supremum of $\mu(A)$ over all sets of finite measure is infinite. Then we can choose a disjoint sequence of sets A_n with each $\mu(A_n)$ finite and with $\sum \mu(A_n) = +\infty$. A

little argument allows us to partition the terms of the series into two subsets, with the series obtained from each subset divergent. Say the terms of one subset are $\mu(B_i)$ and the terms of the other are $\mu(C_j)$. Since $\sum \mu(B_i) = +\infty$, the hypothesis makes $\mu\big(\big(\bigcup_i B_i\big)^c\big)$ finite. A contradiction arises because $\big(\bigcup_i B_i\big)^c \supseteq \bigcup_j C_j$ and $\bigcup_j C_j$ has infinite measure.

8. Consider the set $\mathcal{A}$ of all Borel sets E such that $f^{-1}(E)$ is measurable. The set $\mathcal{A}$ is closed under complements and countable unions, and it contains all intervals. So it is a σ-algebra containing all intervals and must consist of all Borel sets.

10. This problem can be done via dominated convergence, but let us do it from scratch in order to be able to quote it in solving Problem 18 and other problems. We have

$$\big|\int_X f_n\,d\mu - \int_X f\,d\mu\big| \le \int_X |f_n - f|\,d\mu \le \mu(X)\ \sup_x |f_n(x) - f(x)|,$$

and the right side tends to 0 by the uniform convergence. Thus $\lim_X f_n\,d\mu = \int_X f\,d\mu$, the limit existing.

11. In (a) the approximating sets are finite unions of intervals, and we can add their lengths to obtain $\prod_{n=1}^N (1 - r_n)$. Then apply Corollary 5.3. For (b), the set C^c is open, and every point of C^c has an open interval about it where I_C is identically 0; this proves the continuity at points of C^c. To have continuity of I_C at a point x_0 of C, we would need $I_C > 1/2$ on some interval about x_0, and this would mean that I_C equals 1 on that interval and hence that the interval is contained in C. But C contains no intervals of positive length. Part (c) is handled by the same argument as (b). For (d), part (c) says that I_C cannot be redefined on a Lebesgue measurable set of measure 0 so as to be continuous except on a set of measure 0. Theorem 3.29 says that no f obtained by redefining I_C on a set of Lebesgue measure 0 can be Riemann integrable. On the other hand, I_C is measurable, being the indicator function of a compact set, and hence it is Lebesgue integrable.

12. Argue for indicator functions and then simple functions. Then pass to the limit to handle nonnegative functions.

13. Let D be the diagonal. Let $\mathcal{B}$ be the set of all subsets of $\mathcal{A} \times \mathcal{A}$ containing only countably many members of D. This is a σ-ring, and it contains all rectangles of the form $A \times A'$, $A \times B$, and $B \times A$ with A, A', and B^c in $\mathcal{A}$. If $\mathcal{B}_c$ denotes the set of complements of members of $\mathcal{B}$, then Proposition 5.37 shows that $\mathcal{C} = \mathcal{B} \cup \mathcal{B}_c$ is a σ-algebra, and certainly $\mathcal{C}$ contains all rectangles $A^c \times A'^c$ with A and A' in $\mathcal{A}$. Therefore $\mathcal{C} = \mathcal{A} \times \mathcal{A}$. If the set D is in $\mathcal{A} \times \mathcal{A}$, either D or D^c must be in $\mathcal{B}$, and neither is the case.

14. To prove that R is measurable, one first proves the assertion for simple functions ≥ 0 and then passes to the limit. For the rest Fubini's Theorem gives

$$\begin{aligned}\int_{X \times [0,+\infty]} I_R\,d(\mu \times m) &= \int_X \big[\int_{[0,+\infty]} I_R(x, y)\,dm(y)\big]\,d\mu(x)\\ &= \int_X \big[\int_{[0, f(x))} dm(y)\big]\,d\mu(x) = \int_X f(x)\,d\mu(x).\end{aligned}$$

15. This is proved in the same way as Proposition 5.52a.

16. The measure space is the unit interval with Lebesgue measure, and each f_n is an indicator function. The set of which f_n is the indicator function is the subset of $\mathbb{R}$ between $\sum_{k=1}^{n-1} a_k$ and $\sum_{k=1}^{n} a_k$ written modulo 1, i.e., the set of fractional parts of each of these rational numbers. The divergence of the series forces these sets to cycle through the unit interval infinitely often, and thus $f_n(x)$ is 1 infinitely often and 0 infinitely often.

17. From the definition of E_{MN}, we see that $\bigcup_N E_{MN} = X$ and $\bigcap_N E_{MN}^c = \varnothing$. The sets E_{MN} are increasing as a function of N, and their complements are decreasing with empty intersection. Corollary 5.3 produces an integer $C(M)$ such that $\mu(E_{M,C(M)}^c) < \epsilon/2^M$. Put $E = \bigcup_M E_{M,C(M)}^c$. Then $\mu(E) < \epsilon$ by Proposition 5.1g. If $\epsilon' > 0$ is given, we are to produce K such that $|f_k(x) - f(x)| < \epsilon'$ for all $k \geq K$ and all x in E^c. Choose M_0 with $1/M_0 < \epsilon'$. The integer K will be $C(M_0)$. Since x is in $E^c = \bigcap_M E_{M,C(M)}$, x is in $E_{M_0,C(M_0)}$ in particular. Then $|f_k(x) - f(x)| < 1/M_0 < \epsilon'$ for $k \geq C(M_0)$.

18. In (a), we may take the set of integration to be X. Let S be the set of measure 0 on which any of f_n and f is infinite, and redefine all the functions to be 0 on S. Given $\epsilon > 0$, choose $\delta > 0$ by Corollary 5.24 such that $\mu(F) < \delta$ implies $\int_F g\,d\mu < \epsilon$. Let E be as in Egoroff's Theorem for the number δ. Problem 10 shows that $\lim \int_{E^c} f_n\,d\mu = \int_{E^c} f\,d\mu$, the limit existing. Also, $\big|\int_E f_n\,d\mu\big| \leq \int_E |f_n|\,d\mu \leq \int_E g\,d\mu < \epsilon$ for all n, and similarly for f. Hence $\limsup_n \big|\int_X f_n\,d\mu - \int_X f\,d\mu\big| \leq 2\epsilon$. Since ϵ is arbitrary, the result follows.

In (b), consider the measure $g\,d\mu$ and the sequence of functions $\{h_n\}$ with $h_n(x) = f_n(x)/g(x)$ when $g(x) > 0$, $h_n(x) = 0$ when $g(x) = 0$. After checking that h_n is measurable, use Corollary 5.28 and apply (a). The constant that bounds the sequence is 1.

19. By Fatou's Lemma, $\int_{E^c} f\,d\mu \leq \liminf_n \int_{E^c} f_n\,d\mu$. Subtracting this from $\int_X f\,d\mu = \lim \int_X f_n\,d\mu$ gives $\int_E f\,d\mu \geq \limsup_n \int_E f_n\,d\mu$. Another application of Fatou's Lemma gives $\liminf_n \int_E f_n\,d\mu \geq \int_E f\,d\mu$, and we conclude that $\liminf_n \int_E f_n\,d\mu = \limsup_n \int_E f_n\,d\mu = \int_E f\,d\mu$, from which the result follows.

20. Let $\epsilon > 0$ be given. Choose $\delta > 0$ by Corollary 5.24 such that $\mu(F) \leq \delta$ implies $\int_F f\,d\mu \leq \epsilon$. Then choose E with $\mu(E) < \delta$ such that f_n converges to f uniformly off E. Problem 10 shows that there is an N such that $\int_{E^c} |f_n - f|\,d\mu < \epsilon$ for $n \geq N$, and Problem 19 shows that there is an N' such that $\int_E |f_n - f|\,d\mu \leq \int_E f_n\,d\mu + \int_E f\,d\mu \leq 2\int_E f\,d\mu + \epsilon$ for $n \geq N'$. Since $\mu(E) < \delta$, $2\int_E f\,d\mu + \epsilon \leq 3\epsilon$. Then $n \geq \max\{N, N'\}$ implies $\int_X |f_n - f|\,d\mu \leq 4\epsilon$.

21. Suppose that $\lim \int_X f_n\,d\mu = \int_X f\,d\mu$. Given $\epsilon > 0$, choose $\delta > 0$ by Corollary 5.24 such that $\mu(E) < \delta$ implies $\int_E f\,d\mu < \epsilon$. Then choose N such that $N^{-1}\big(\int_X f\,d\mu + \epsilon\big) < \delta$. For any n, the convergence of $\int_X f_n\,d\mu$ to $\int_X f\,d\mu$ implies that $N\mu(\{x \mid f_n(x) \geq N\}) \leq \int_{\{x \mid f_n(x) \geq N\}} f_n\,d\mu \leq \int_X f_n\,d\mu \leq \int_X f\,d\mu + \epsilon$ if n is sufficiently large. Hence $\mu(\{x \mid f_n(x) \geq N\}) \leq N^{-1}\big(\int_X f\,d\mu + \epsilon\big) < \delta$ for large

n, and therefore $\int_{\{x \mid f_n(x) \geq N\}} f\,d\mu < \epsilon$. Problem 20 shows that $\int_X |f_n - f|\,d\mu \leq \epsilon$ if n is large enough, and then also $\int_{\{x \mid f_n(x) \geq N\}} |f_n - f|\,d\mu \leq \epsilon$. So we have $\int_{\{x \mid f_n(x) \geq N\}} f_n\,d\mu \leq \int_{\{x \mid f_n(x) \geq N\}} |f_n - f|\,d\mu + \int_{\{x \mid f_n(x) \geq N\}} f\,d\mu \leq \epsilon + \epsilon = 2\epsilon$ for n large, say $n \geq N'$. By increasing N and taking the integrability of $f_1, \dots, f_{N'-1}$ into account, we can achieve the inequality $\int_{\{x \mid f_n(x) \geq N\}} f_n\,d\mu \leq 2\epsilon$ for all n.

Conversely suppose that $\{f_n\}$ is uniformly integrable. Given $\epsilon > 0$, find the N of uniform integrability, put $\delta = \epsilon/N$, and choose E_0 by Egoroff's Theorem such that $\mu(E_0) < \delta$ and f_n converges uniformly off E_0. Then $\lim \int_{E_0^c} f_n\,d\mu = \int_{E_0^c} f\,d\mu$ by Problem 10. Fatou's Lemma gives $\int_{E_0} f\,d\mu \leq \liminf \int_{E_0} f_n\,d\mu$, and we have

$$\int_{E_0} f_n\,d\mu = \int_{E_0 \cap \{x \mid f_n(x) \geq N\}} f_n\,d\mu + \int_{E_0 - \{x \mid f_n(x) \geq N\}} f_n\,d\mu.$$

The first term on the right side is $\leq \int_{\{x \mid f_n(x) \geq N\}} f_n\,d\mu$, which is $\leq \epsilon$ by uniform integrability, and the second term on the right side is $\leq N\delta = \epsilon$ because $\mu(E_0) < \delta$ and $f_n(x) \leq N$ on the set of integration. Thus $\limsup \int_{E_0} f_n\,d\mu \leq 2\epsilon$, and we obtain $\limsup_n \left| \int_{E_0} f_n\,d\mu - \int_{E_0} f\,d\mu \right| \leq 4\epsilon$.

22. In the notation of Section 5, $\mathcal{K} = \mathcal{U} = \mathcal{A}$ since $\mathcal{A}$ is now assumed to be a σ-algebra. Thus $\mu_*(E) = \sup_{K \in \mathcal{A},\, K \subseteq E} \mu(K)$ and $\mu^*(E) = \inf_{U \in \mathcal{A},\, U \supseteq E} \mu(U)$. Take a sequence of sets K_n in $\mathcal{A}$ with $\lim \mu(K_n) = \mu_*(E)$; without loss of generality, the sets K_n may be assumed increasing. Then we may take K to be the union of the K_n. The construction of U is similar.

The set K is any member of $\mathcal{A}$ such that $\mu(K)$ is the supremum of $\mu(S)$ for all S in $\mathcal{A}$ with $S \subseteq E$. Then $\mu(K^c)$ is the infimum of all $\mu(S^c) = \mu(X) - \mu(S)$ for all S^c in $\mathcal{A}$ with $S^c \supseteq E^c$. A similar argument applies to U and U^c. The result is that $U^c \subseteq E^c \subseteq K^c$, $\mu_*(E^c) = \mu(U^c)$, and $\mu^*(E^c) = \mu(K^c)$.

23. Lemma 5.33 gives $\mu(A \cap K) \leq \mu_*(A \cap E)$, $\mu(A^c \cap K) \leq \mu_*(A^c \cap E)$, and $\mu_*(E) = \mu(K) = \mu(A \cap K) + \mu(A^c \cap K) \leq \mu_*(A \cap E) + \mu_*(A^c \cap E) \leq \mu_*(E)$, from which we obtain $\mu_*(A \cap E) = \mu(A \cap K)$. The argument that $\mu^*(A \cap E) = \mu(A \cap U)$ is similar.

24. The right side of the definition of σ depends only on $A \cap E$ and $B \cap E^c$, and hence σ is well defined. The formulas

$$\bigcup_n \left[(A_n \cap E) \cup (B_n \cap E^c)\right] = \Big(\Big(\bigcup_n A_n\Big) \cap E\Big) \cup \Big(\Big(\bigcup_n B_n\Big) \cap E^c\Big)$$

and $[(A \cap E) \cup (B \cap E^c)]^c = (A^c \cap E) \cup (B^c \cap E^c)$ show that the sets in question form a σ-algebra $\mathcal{C}$. Taking $A = B$ shows that $\mathcal{A} \subseteq \mathcal{C}$, and taking $A = X$ and $B = \varnothing$ shows that E is in $\mathcal{C}$. Therefore $\mathcal{B} \subseteq \mathcal{C}$, and σ is defined on all of $\mathcal{B}$.

The complete additivity of σ results from the complete additivity of each of the four terms in the definition of σ. Specifically let a disjoint sequence $(A_n \cap E) \cup (B_n \cap E)$ be given, and let $A = \bigcup_n A_n$ and $B = \bigcup_n B_n$. We have $\mu_*(A_n \cap E) = \mu(A_n \cap K)$, and the sets $A_n \cap K$ are disjoint; thus $\sum \mu_*(A_n \cap E) = \mu_*(A \cap E)$. The next term is $\mu^*(A_n \cap E) = \mu(A_n \cap U)$, and the sets $A_n \cap U$ may not be disjoint. However, $\mu^*(A_m \cap E) + \mu^*(A_n \cap E) = \mu(A_m \cap U) + \mu(A_n \cap U) = \mu(A_m \cap A_n \cap E) + \mu((A_1 \cup A_2) \cap E)$, and $\mu(A_m \cap A_n \cap U) = \mu^*(A_m \cap A_n \cap E) = \mu^*(\varnothing) = 0$. Thus the term with $\mu^*(A_n \cap E)$ behaves in additive fashion. Consequently $\mu^*(A \cap E) \geq \mu^*\big(\big(\bigcup_{k=1}^n A_k\big) \cap E\big) = \sum_{k=1}^n \mu^*(A_k \cap E)$. Letting n tend to infinity gives $\mu^*(A \cap E) \geq \sum_{k=1}^\infty \mu^*(A_k \cap E)$. The reverse inequality follows from Lemma 5.33a, and thus the term $\mu^*(A_n \cap E)$ is completely additive. The terms with the B_n's are handled similarly, and σ is completely additive.

Taking $A = X$ and $B = \varnothing$, we see immediately that the formula for $\sigma(E)$ is as asserted.

To prove that $\sigma(A) = \mu(A)$ for A in $\mathcal{A}$, we take $A = B$. Then we see that $\sigma(A) = t\mu(A \cap K) + (1-t)\mu(A \cap U) + t\mu(A \cap K^c) + (1-t)\mu(A \cap U^c) = t\mu(A) + (1-t)\mu(A) = \mu(A)$.

25. Each member of the countable set has only countably many ordinals less than it, and the countable union of countable sets is countable. Therefore some member of Ω is not accounted for and is an upper bound for the countable set. Application of (iii) completes the argument.

27. For (a), if $U_n \uparrow U$ and $V_n \uparrow V$, then $U_n \cup V_n \uparrow U \cup V$ and $U_n \cap V_n \uparrow U \cap V$. Similar remarks apply to $\mathcal{K}_\alpha$. Then the assertion follows by transfinite induction.

For (b), we know that $\mathcal{K}_\alpha$ is closed under finite unions and intersections, and we readily see that the complement of any set occurs at most one step later. Now let an increasing sequence of sets in various $\mathcal{K}_\alpha$'s be given. Say that U_n is in $\mathcal{K}_{\alpha_n}$. Problem 25 shows that there is a countable ordinal α_0 that is $\geq$ all the α_n, and then all the U_n are in $\mathcal{K}_{\alpha_0}$. The union is then in $\mathcal{U}_{\alpha_0+1}$ and necessarily in $\mathcal{K}_{\alpha_0+1}$. Hence the union is in the union of the $\mathcal{K}_\alpha$'s. So the union of the $\mathcal{K}_\alpha$'s is a σ-algebra and must contain $\mathcal{B}$. All the set-theoretic operations take place within $\mathcal{B}$, and thus the union must actually equal $\mathcal{B}$.

28. Proposition 5.2 and Corollary 5.3 show that the value of the measure is determined on all the new sets that are constructed in terms of the values on the previous sets. Problem 27 shows that all members of $\mathcal{B}$ are obtained by the construction, and hence μ is completely determined on $\mathcal{B}$.

29. Same argument as for Problem 27b.

30. At every stage of taking limits, we have closure under addition and scalar multiplication. Pointwise decreasing limits produce the indicator functions of finite unions of closed intervals, and pointwise increasing limits of them produce the indicator functions of arbitrary finite unions of intervals. Since the constants are present as continuous functions, we have the indicator function of every elementary set and its complement. These sets form an algebra. Going through the construction of Problem 27, we obtain the indicator function of every Borel set. Since we have closure under addition and scalar multiplication at each step, we obtain all simple functions. One increasing limit gives us all nonnegative Borel measurable functions,

and a subtraction (allowable without another passage to the limit) gives us all Borel measurable functions.

32. To see that C has the same cardinality as $\mathbb{R}$, we can make an identification of the disjoint union of $\mathbb{R}$ and a countable set. To do so, we write C as the members of $[0, 1]$ whose base-3 expansions involve no 1's. For each such infinite sequence of 0's and 2's, we change all the 2's to 1's and regard the result as the base-2 expansion of some real number. This identification is onto $[0, 1]$, and it is one-one if we discard from C all the sequences of 0's and 2's that end in all 2's.

The standard Cantor set has Lebesgue measure 0, and thus any subset of it is Lebesgue measurable of measure 0. The cardinality of this set of subsets is the same as the cardinality of the set of subsets of $\mathbb{R}$. In Section A.10 of the appendix, it is shown for any set S that the cardinality of S is less than the cardinality of the set of all subsets of S. So the cardinality of the set of Lebesgue measurable sets is at least that of the set of all subsets of $\mathbb{R}$.

33. Since C^c is open, any member x of C^c has the property that $I_{C'}$ is 0 on some open interval about x. Thus $I_{C'}$ is continuous at x. Since C has Lebesgue measure 0, $I_{C'}$ is continuous except on a Lebesgue measurable set of measure 0. Theorem 3.29 shows that $I_{C'}$ is Riemann integrable. Hence the cardinality of the set of Riemann integrable functions is at least that of the set of all subsets of $\mathbb{R}$.

35. If $\mathcal{F}$ is the given filter, form the partially ordered set consisting of all filters on X containing $\mathcal{F}$, with inclusion as the partial ordering. The union of the members of a chain is readily verified to be an upper bound for the chain, and Zorn's Lemma produces a maximal element. This maximal element is readily seen to be an ultrafilter.

36. The filter in question consists of all supersets of finite intersections of members of $\mathcal{C}$.

37–38. Suppose that $\mathcal{F}$ is an ultrafilter, $A \cup B$ is in $\mathcal{F}$, A is not in $\mathcal{F}$, and B is not in $\mathcal{F}$. Let $\mathcal{F}'$ consist of all sets in $\mathcal{F}$ and all sets $B \cap F$ with F in $\mathcal{F}$. Since B is not in $\mathcal{F}$, $\mathcal{F}'$ properly contains $\mathcal{F}$. Since $\mathcal{F}$ is an ultrafilter, $\mathcal{F}'$ must fail to be a filter. On the other hand, by inspection, $\mathcal{F}'$ satisfies properties (i) and (ii) in the definition of filter. We conclude that $\varnothing$ is in $\mathcal{F}'$, hence that there is a set F in $\mathcal{F}$ with $B \cap F = \varnothing$. Since $\mathcal{F}$ satisfies (ii), the set $(A \cup B) \cap F = (A \cap F) \cup (B \cap F) = A \cap F$ is in $\mathcal{F}$. By (i), A is in $\mathcal{F}$, contradiction.

Conversely suppose that $\mathcal{F}$ is a filter such that either A or A^c is in $\mathcal{F}$ for each subset A of X. If $\mathcal{F}$ is not maximal, let B be a set that lies in some filter $\mathcal{F}'$ properly containing $\mathcal{F}$ while B is not itself in $\mathcal{F}$. By hypothesis, B^c is in $\mathcal{F}$ and hence is in $\mathcal{F}'$. But then $B \cap B^c = \varnothing$ lies in $\mathcal{F}'$, in contradiction to (iii).

39. If an ultrafilter $\mathcal{F}$ is given, define $\mu(E) = 1$ if E is in $\mathcal{F}$ and define $\mu(E) = 0$ otherwise. Then μ is defined on all subsets, and we have $\mu(\varnothing) = 0$ and $\mu(X) = 1$. If E and E' are disjoint, we are to show that

$$\mu(E) + \mu(E') = \mu(E \cup E').$$

If $E \cup E'$ is not in $\mathcal{F}$, then all terms in the displayed equation are 0 since $\mathcal{F}$ is closed under supersets. If $E \cup E'$ is in $\mathcal{F}$, then Problem 37 shows that E or E' is in $\mathcal{F}$; on the other hand, they cannot both be in $\mathcal{F}$ because $\mathcal{F}$ is closed under finite intersections and the empty set is not in $\mathcal{F}$. Thus exactly one term on the left side of the displayed equation is 1, and the right side is 1. This proves additivity.

Conversely if an additive set function μ is given on all subsets of X that takes only the values 0 and 1 and is not the 0 set function, let $\mathcal{F}$ consist of the sets E for which $\mu(E) = 1$. It is immediate that (i) and (iii) hold in the definition of filter. For (ii), let E and E' be in $\mathcal{F}$. Then $E \cup E'$ is in $\mathcal{F}$. Hence $\mu(E \cap E') + 1 = \mu(E) + \mu(E') = 1 + 1$, and $\mu(E \cap E') = 1$. Hence $\mathcal{F}$ is closed under finite intersections and (ii) holds. Thus $\mathcal{F}$ is a filter. If A is given, we have $1 = \mu(X) = \mu(A) + \mu(A^c)$, and hence exactly one of the sets A and A^c is in $\mathcal{F}$. By Problem 38, $\mathcal{F}$ is an ultrafilter.

The statement that complete additivity is equivalent to closure of the ultrafilter under countable intersections is a routine consequence of Corollary 5.3.

40. This follows from Problems 34d and 35.

41. Let S_n be the set of all integers $\geq n$. Since $S_1 = X$, S_1 is in the ultrafilter. Since the ultrafilter is not trivial, $\{n\}$ is not in it, and thus Problem 37 shows that S_n is in it if S_{n-1} is in it. Hence S_n is in the ultrafilter for all n. The countable intersection $\bigcap_n S_n$ is empty, and the empty set is not in any filter. Hence the ultrafilter is not closed under countable intersections. Corollary 5.3 shows that the corresponding set function is not completely additive.

43. The proof of Proposition 5.26 shows that the result holds for simple functions ≥ 0. If $f \geq 0$ and $g \geq 0$, choose the standard sequences t_n and u_n of simple functions increasing to f and g. These converge uniformly. Hence so does the sum $s_n = t_n + u_n$. The same argument as for Problem 10 shows that $\lim \int_E s_n \, d\mu = \int_E (f + g) \, d\mu$, $\lim \int_E t_n \, d\mu = \int_E f \, d\mu$, and $\lim \int_E u_n \, d\mu = \int_E g \, d\mu$. Thus the result holds for bounded nonnegative f and g. The passage to general bounded f and g is achieved as in Proposition 5.26.

Chapter VI

1. In additive notation, the sets $E + t$ for t in T are disjoint, and their countable union is S^1. Since Lebesgue measure is translation invariant, these sets all have the same measure c. Then complete additivity gives $c\,\infty = 2\pi$, which is impossible.

2. Parts (b) and (c) are easy. For (a), expand the Jacobian determinant $J(N)$ in cofactors about the first row, obtaining two terms—one each from the first two entries of the first row. The first term is $\cos\theta_1$ times a determinant of size $N - 1$ whose first column has a common factor of $r\cos\theta_1$ and whose second column has a common factor of $\sin\theta_1$, the remaining part of the determinant being $J(N - 1)$; thus the first term gives $(r\cos^2\theta_1 \sin\theta_1) J(N - 1)$. The second term is $-(-r\sin\theta_1)$ times a determinant of size $N - 1$ whose first column has a common factor of $\sin\theta_1$

and whose second column has a common factor of $r\sin\theta_1$, the remaining part of the determinant being $J(N-1)$; thus the second term gives $(r\sin^3\theta_1)J(N-1)$. Adding the two terms gives $J(N)=(r\sin\theta_1)J(N-1)$, and an induction readily proves the formula.

3. Replace f in Theorem 6.32 by $f\circ L$, and use $\varphi=L^{-1}$. Since $\varphi'(x)=L^{-1}$ for each x, the result follows.

4. In the result of Problem 3, use $L(x)=yx$ and replace $f(z)$ by $f(z)/|\det z|^N$. Then the left side in Problem 3 is $\int_{M_N} f(yx)/|\det(yx)|^N\,dx$, while the right side is $|\det L|^{-1}\int_{M_N} f(x)/|\det x|^N\,dx$. Thus $|\det y|^{-N|}\int_{M_N} f(yx)/|\det(x)|^N\,dx = |\det L|^{-1}\int_{M_N} f(x)/|\det x|^N\,dx$, and the problem reduces to showing that $\det L=(\det y)^N$. One way of doing this is to verify that this formula is true if y is the matrix of an elementary row operation and then to multiply the results. But a faster way is to let $x_1,\dots,x_n$ be the columns of x, so that $L(x_1,\dots,x_n)=(yx_1,\dots,yx_n)$. Then L as a matrix is given in block diagonal form by a copy of y in each block. Hence $\det L=(\det y)^n$. In a little more detail, the matrix of L is being formed relative to the following basis of M_N: if E_{ij} is the N-by-N matrix with 1 in the $(i,j)^{\text{th}}$ entry and 0 elsewhere, the basis is $E_{11}, E_{21},\dots,E_{N1},E_{12},\dots,E_{NN}$.

5. For (a), we have, for $n\neq 0$,

$$2\pi c_n=\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx=\int_{|x|\le\frac{1}{|n|}} f(x)e^{-inx}\,dx+\int_{\frac{1}{|n|}\le|x|\le\pi} f(x)e^{-inx}\,dx.$$

Let us call these terms I and II. Since $|f(x)|\le C|x|^\alpha$ for $|x|\le 1$,

$$|I|\le\int_{|x|\le\frac{1}{|n|}}|f(x)|\,dx\le C\int_{|x|\le\frac{1}{|n|}}|x|^\alpha\,dx=\tfrac{2C}{1+\alpha}\,\tfrac{1}{|n|^{1+\alpha}}.$$

For II, we use integration by parts and take into account that the terms at π and $-\pi$ cancel by periodicity:

$$\begin{aligned}II&=\Big(\int_{-\pi}^{-1/|n|}+\int_{1/|n|}^{\pi}\Big)f(x)\,dx\\&=\Big[\tfrac{f(x)e^{-inx}}{-in}\Big]_{-\pi}^{-1/|n|}+\Big[\tfrac{f(x)e^{-inx}}{-in}\Big]_{1/|n|}^{\pi}+\tfrac{1}{in}\int_{\frac{1}{|n|}\le|x|\le\pi} f'(x)e^{-inx}\,dx\\&=\tfrac{1}{in}\Big\{f\big(\tfrac1n\big)e^{-in/|n|}-f\big(-\tfrac1n\big)e^{+in/|n|}\Big\}+\tfrac{1}{in}\int_{\frac{1}{|n|}\le|x|\le\pi} f'(x)e^{-inx}\,dx.\end{aligned}$$

Let us call the terms on the right III and IV. Since $|f(x)|\le C|x|^\alpha$ for $|x|\le 1$,

$$|III|\le\tfrac{1}{|n|}\Big(\big|f\big(\tfrac1n\big)\big|+\big|f\big(-\tfrac1n\big)\big|\Big)\le 2C\,\tfrac{1}{|n|^{1+\alpha}}.$$

The derivation of the formula for II, when applied to f' instead of f, gives the following value for IV:

$$IV=-\tfrac{1}{n^2}\Big\{f'\big(\tfrac1n\big)e^{-in/|n|}-f'\big(-\tfrac1n\big)e^{+in/|n|}\Big\}-\tfrac{1}{n^2}\int_{\frac{1}{|n|}\le|x|\le\pi} f''(x)e^{-inx}\,dx.$$

Let us call the terms on the right V and VI. Since $|f'(x)| \le C|x|^{\alpha-1}$ for $|x| \le 1$,

$$|V| \le \tfrac{1}{n^2}\big(\big|f'\big(\tfrac{1}{n}\big)\big| + \big|f'\big(-\tfrac{1}{n}\big)\big|\big) \le 2C\,\tfrac{1}{|n|^{1+\alpha}}.$$

Since $f''(x)$ is bounded for $1 \le |x| \le \pi$, we can write $|f''(x)| \le C'|x|^{\alpha-2}$ for $0 < |x| \le \pi$, in view of the assumption on f''. Therefore

$$\begin{aligned} VI &\le \tfrac{1}{n^2}\int_{\frac{1}{|n|}\le|x|\le\pi} C'|x|^{\alpha-2}\,dx = \tfrac{2C'}{n^2}\int_{1/|n|}^{\pi} x^{\alpha-2}\,dx \\ &= \tfrac{2C'}{1-\alpha}\,\tfrac{1}{n^2}\big(\tfrac{1}{|n|^{\alpha-1}} - \pi^{\alpha-1}\big) \le \tfrac{2C'}{1-\alpha}\,\tfrac{1}{|n|^{1+\alpha}}. \end{aligned}$$

Since $2\pi|c_n| \le |I| + |III| + |V| + |VI|$, we obtain $|c_n| \le K/|n|^{1+\alpha}$.

For (b), the uniform convergence follows by applying the Weierstrass M-test, and the limit is f as a consequence of the uniqueness theorem.

In (c), a proof is called for. The crux of the matter is to show, under the assumption that f is real valued, that the variation V_ε of f on $[\varepsilon, 1]$, which gets larger as ε decreases to 0, is bounded. If $x_0 < \cdots < x_n$ is a partition P of $[\varepsilon, 1]$, then

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| = \sum_{i=1}^{n} |f'(\xi_i)|(x_i - x_{i-1}) \le C\sum_{i=1}^{n} \xi_i^{\alpha-1}(x_i - x_{i-1})$$

with $x_{i-1} < \xi_i < x_i$. With ε fixed, the right side is a Riemann sum for the bounded function $x^{\alpha-1}$ on $[\varepsilon, 1]$ and is $\le$ the corresponding upper sum $U(P, x^{\alpha-1}\big|_{[\varepsilon,1]})$. As we insert points into the partition, the left sides increase and the right sides decrease to the limit $\int_\varepsilon^1 x^{\alpha-1}\,dx = \alpha^{-1}(1-\varepsilon^\alpha)$. Hence $V_\varepsilon \le C\alpha^{-1}(1-\varepsilon^\alpha)$, and $\sup_{\varepsilon>0} V_\varepsilon \le C/\alpha$.

6. The distribution function F of μ must have $F(b) - F(a)$ equal to 0 or 1 for all a and b. If c is the supremum of the x's for which there exists $y > x$ with $F(x) < F(y)$, then F has to be k on $(-\infty, c)$ and $k+1$ on $[c, +\infty)$ for the value of k that makes $F(0) = 0$. Hence μ is a point mass at c with $\mu(\{c\}) = 1$.

7. Let K be compact, and let f and g both be equal to the members of a sequence $\{f_n\}$ of continuous functions of compact support decreasing to the indicator function I_K of K. Applying the identity to f_n and passing to the limit, we obtain $\nu(K) = \nu(K)^2$. Thus $\nu(K)$ is 0 or 1 for each compact set. By regularity ν takes on only the values 0 and 1 on Borel sets. Then the argument (but not the statement) of Problem 6 applies, and there is some c with ν equal to a point mass at c with $\nu(\{c\}) = 1$.

8. In (a), if the complement of the set in question is not dense, it omits an open set. However, nonempty open sets have positive measure.

In (b), form $\int_{\mathbb{R}^1}\big[\int_{\mathbb{R}^1} I_E(x-t)\,dt\big]\,d\mu(x)$. The inner integral equals the Lebesgue measure of E for every x since Lebesgue measure is invariant under translations and the map $t \mapsto -t$. Hence the iterated integral is 0. The integral in the other order is $0 = \int_{\mathbb{R}^1}\big[\int_{\mathbb{R}_1} I_E(x-t)\,d\mu(x)\big]\,dt = \int_{\mathbb{R}^1}\big[\int_{\mathbb{R}_1} I_{E+t}(x)\,d\mu(x)\big]\,dt = \int_{\mathbb{R}^1}\mu(E+t)\,dt$, and Corollary 5.23 shows that $\mu(E+t)$ is 0 almost everywhere.

In (c), the same computation applies, and $\mu(E+t)$ is 0 almost everywhere. Under the assumption that $\lim_{t\to 0}\mu(E+t)$ exists, the limit must be 0, by (a).

9. Write $1/|x|$ as a sum $F_1 + F_\infty$, where F_1 is $1/|x|$ for $|x| < 1$ and is 0 for $|x| \geq 1$. Then $\int_{\mathbb{R}^3} F_\infty(x-y)\,d\mu(y)$ is bounded by $\mu(\mathbb{R}^3)$, and it is enough to handle the contribution from F_1. For that we have $\int_{\mathbb{R}^3}\big[\int_{\mathbb{R}^3} F_1(x-y)\,d\mu(y)\big]\,dx = \int_{\mathbb{R}^3}\big[\int_{\mathbb{R}^3} F_1(x-y)\,dx\big]\,d\mu(y) = \int_{\mathbb{R}^3}\big[\int_{\mathbb{R}^3} F_1(x)\,dx\big]\,d\mu(y) = \mu(\mathbb{R}^3)\int_{|x|\leq 1}|x|^{-1}\,dx$, and this is finite in $\mathbb{R}^3$. Hence the inner integral $\int_{\mathbb{R}^3} F_1(x-y)\,d\mu(y)$ is finite almost everywhere.

10. We proceed by induction on n, the case $n = 1$ following since finite sets have Lebesgue measure 0. Assume the result in $n-1$ variables, and let $P(x_1,\dots,x_n) \not\equiv 0$ be given. Let E be the set where $P = 0$. This is closed, hence Borel measurable in $\mathbb{R}^n$. Fix $(x_1',\dots,x_n')$ with $P(x_1',\dots,x_n') \neq 0$. The polynomial in one variable $R(x) = P(x_1',\dots,x_{n-1}',x)$ is not identically 0, being nonzero at $x = x_n'$, and hence it vanishes only finitely often, say for x in the finite set F. Fix $x' \notin F$. Then the polynomial $Q(x_1,\dots,x_{n-1}) = P(x_1,\dots,x_{n-1},x')$ in $n-1$ variables is not identically 0, being nonzero at $(x_1',\dots,x_{n-1}')$, and its set $E_{x'}$ of zeros has measure 0 by inductive hypothesis. If m_n denotes n-dimensional Lebesgue measure, then Fubini's Theorem applied to I_E gives

$$m_n(E) = \int_{\mathbb{R}} m_{n-1}(E_{x'})\,dx = \int_F m_{n-1}(E_{x'})\,dx' + \int_{F^c} m_{n-1}(E_{x'})\,dx'.$$

On the right side the first term is 0 since the 1-dimensional measure of F is 0, while the second term is 0 since the integrand is 0. Thus $m(E) = 0$.

11. $\Gamma(x+y)\int_0^1 t^{x-1}(1-t)^{y-1}\,dt = \int_0^\infty e^{-s}s^{x+y-1}\,ds\int_0^1 t^{x-1}(1-t)^{y-1}\,dt = \int_0^\infty\big[\int_0^s u^{x-1}(s-u)^{y-1}e^{-s}\,du\big]\,ds = \int_0^\infty\big[\int_u^\infty u^{x-1}(s-u)^{y-1}e^{-s}\,ds\big]\,du = \int_0^\infty\big[\int_0^\infty u^{x-1}s^{y-1}e^{-s}e^{-u}\,ds\big]\,du = \Gamma(x)\Gamma(y)$.

12. In Cartesian coordinates we obtain 1^N, hence 1. In spherical coordinates we obtain $\Omega_{N-1}\int_0^\infty r^{N-1}e^{-\pi r^2}\,dr$. Putting $\pi r^2 = s$ shows that $\int_0^\infty r^{N-1}e^{-\pi r^2}\,dr = \int_0^\infty (s/\pi)^{(N-2)/2}e^{-s}\frac{1}{2\pi}\,ds = \frac{1}{2}\pi^{-N/2}\Gamma(N/2)$. Hence $\Omega_{N-1} = 2\pi^{N/2}/\Gamma(N/2)$.

13. Part (a) is carried out by showing by induction on k that $\sum_{i=1}^k x_i = 1 - \prod_{i=1}^k(1-u_i)$. The case $k = n$ is the desired result.

In (b), let $0 < u_i < 1$ for all i. Then $x_i > 0$ for all i, and (a) makes it clear that $\sum_{i=1}^n x_i < 1$. Therefore φ carries I into S. Define $u = \widetilde{\varphi}(x)$ by the formula in (b). If all $x_i > 0$ and $\sum_{i=1}^n x_i < 1$, then certainly $u_i > 0$. Also, $\sum_{j=1}^i x_i < 1$ implies $x_i < 1 - \sum_{j=1}^{i-1} x_j$, so that $u_i = x_i\big/\big(1 - \sum_{j=1}^{i-1} x_j\big) < 1$. Therefore $\widetilde{\varphi}$ carries S into I. To complete the proof, we show that $\widetilde{\varphi}\circ\varphi$ is the identity on I and $\varphi\circ\widetilde{\varphi}$ is the identity on S. For $\widetilde{\varphi}\circ\varphi$, we pass from u to x to v. Thus we start with v_i, substitute the x's, use the inductive version of (a) to substitute the u's, and then sort matters out to see that $v_i = u_i$. For $\varphi\circ\widetilde{\varphi}$, we pass from x to u to y. Then we start with y_i and substitute the u's to obtain $y_i = \big(\prod_{l=1}^{i-1}(1-u_i)\big)u_i$. To substitute for the u's in terms of the x's, we use the inductive version of (a) in the form $\sum_{l=1}^{i-1} y_l = 1 - \prod_{l=1}^{i-1}(1-u_l)$. This gives $\big(\prod_{l=1}^{i-1}(1-u_i)\big)u_i = \big(1 - \sum_{l=1}^{i-1} y_l\big)x_i\big/\big(1 - \sum_{l=1}^{i-1} x_l\big)$. Then an induction on i shows that $y_i = x_i$, and hence $\varphi\circ\widetilde{\varphi}$ is the identity on S.

In (c), routine computation shows that $\varphi'(u)$ is lower triangular with diagonal entries $1, (1-u_1), (1-u_1)(1-u_2), \dots, (1-u_1)\cdots(1-u_{n-1})$, and hence the determinant is the product of these diagonal entries. Similarly $\widetilde{\varphi}'(x)$ is lower triangular with diagonal entries $1, (1-x_1)^{-1}, (1-x_1-x_2)^{-1}, \dots, (1-x_1-x_2-\cdots-x_{n-1})^{-1}$, and its determinant is the product of these diagonal entries.

14. The change of variables in Problem 13 gives

$$\begin{aligned}\int_S x_1^{a_1-1}\cdots x_n^{a_n-1}\,dx &= \int_I u_1^{a_1-1}[(1-u_1)u_2]^{a_2-1}\cdots[(1-u_1)\cdots(1-u_{n-1})u_n]^{a_n-1}\\ &\qquad\times(1-u_1)^{n-1}\cdots(1-u_{n-1})\,du\\ &= \int_I u_1^{a_1-1}(1-u_1)^{a_2+\cdots+a_n-(n-1)+(n-1)}u_2^{a_2-1}\\ &\qquad\times(1-u_2)^{a_3+\cdots+a_n-(n-2)+(n-2)}\\ &\qquad\times\cdots\times u_{n-1}^{a_{n-1}-1}(1-u_{n-1})^{a_n-1+1}u_n^{a_n-1}\,du\\ &= \int_0^1 u_1^{a_1-1}(1-u_1)^{a_2+\cdots+a_n}\,du_1\cdot\int_0^1 u_2^{a_2-1}(1-u_2)^{a_3+\cdots+a_n}\,du_2\\ &\qquad\cdot\dots\cdot\int_0^1 u_{n-1}^{a_{n-1}-1}(1-u_{n-1})^{a_n}\,du_{n-1}\cdot\int_0^1 u_n^{a_n-1}\,du_n.\end{aligned}$$

The right side is the product of 1-dimensional integrals of the kind treated in Problem 11. Substitution of the values from that problem leads to the desired result.

15. The monotonicity makes possible the estimate of uniform convergence, and the continuity then makes the limit continuous. A continuous function is determined by its values on a dense set, and C^c is dense.

16. For each n, $F_n(x) = 1-F_n(1-x)$. Thus $\int_0^1 F_n(x)\,dx = 1-\int_0^1 F_n(1-x)\,dx = 1-\int_0^1 F_n(x)\,dx$ and $\int_0^1 F_n(x)\,dx = \frac{1}{2}$. Passing to the limit and using uniform or dominated convergence, we obtain $\int_0^1 F(x)\,dx = \frac{1}{2}$.

18. Use Proposition 6.47. Then u is harmonic by Problem 14 at the end of Chapter III.

19. Since P_r has L^1 norm 1, the inequality $\|u(r,\,\cdot\,)\|_p \le \|f\|_p$ follows from Minkowski's inequality for integrals. For the limiting behavior as r increases to 1, we extend f periodically and write

$$\begin{aligned}u(r,\theta)-f(\theta) &= \tfrac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\varphi)f(\theta-\varphi)\,d\varphi - f(\theta)\\ &= \tfrac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\varphi)[f(\theta-\varphi)-f(\theta)]\,d\varphi,\end{aligned}$$

the second step following since $\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r\,d\varphi = 1$. Applying Minkowski's inequality for integrals, we obtain

$$\|u(r,\,\cdot\,)-f\|_p \le \tfrac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\varphi)\|f(\theta-\varphi)-f(\theta)\|_{p,\theta}$$

since $P_r \geq 0$. The integration on the right is broken into two sets, $S_1 = (-\delta, \delta)$ and $S_2 = [-\pi, -\delta] \cup [\delta, \pi]$, and the integral is

$$\begin{aligned} &\leq \tfrac{1}{2\pi} \textstyle\int_{S_1} P_r(\varphi)\big(\sup_{\varphi \in S_1} \|f(\theta - \varphi) - f(\theta)\|_{p,\theta}\big)\, d\varphi + \tfrac{1}{2\pi} \textstyle\int_{S_2} P_r(\varphi) 2\|f\|_p \, d\varphi \\ &\leq \sup_{\varphi \in S_1} \|f(\theta - \varphi) - f(\theta)\|_{p,\theta} + 2\|f\|_p \sup_{\varphi \in S_2} P_r(\varphi). \end{aligned}$$

Let $\epsilon > 0$ be given. If δ is sufficiently small, Proposition 6.16 shows that the first term is $< \epsilon$. With δ fixed, we can then choose r close enough to 1 to make the second term $< \epsilon$.

20. For (a), we argue as in Problem 19, taking S_1 and S_2 to be as in that solution. Then

$$\begin{aligned} |u(r, \theta) - f(\theta)| &\leq \tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} P_r(\varphi) |f(\theta - \varphi) - f(\theta)| \, d\varphi \\ &\leq \tfrac{1}{2\pi} \textstyle\int_{S_1} P_r(\varphi) |f(\theta - \varphi) - f(\theta)| \, d\varphi \\ &\quad + \tfrac{1}{2\pi} \textstyle\int_{S_2} P_r(\varphi) [\|f\|_\infty + \sup_{\theta \in E} |f(\theta)|] \, d\varphi \\ &\leq \sup_{\varphi \in S_1} |f(\theta - \varphi) - f(\theta)| \\ &\quad + \big(\sup_{\varphi \in S_2} P_r(\varphi)\big) [\|f\|_\infty + \sup_{\theta \in E} |f(\theta)|], \end{aligned}$$

and the uniform convergence follows.

For (b), the Poisson integral of f is of the form $\sum_{n=-\infty}^{\infty} c_n r^{|n|} e^{in\theta}$, where the c_n are the Fourier coefficients of f. Any other harmonic function in the disk is of the form $\sum_{n=-\infty}^{\infty} c'_n r^{|n|} e^{in\theta}$. Suppose this tends uniformly to f as r increases to 1. Then the difference is a series $\sum_{n=-\infty}^{\infty} d_n r^{|n|} e^{in\theta}$ that converges uniformly to 0. Then the integral of the product of this series and $e^{-ik\theta}$ tends to 0. Interchanging integral and sum, we see that $d_k r^{|k|}$ tends to 0 for each k. Therefore $d_k = 0$ for each k.

In (c) since P_r is even,

$$\begin{aligned} \textstyle\int_{-\pi}^{\pi} (P_r * f)(\theta) g(\theta) \, d\theta &= \textstyle\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} P_r(\theta - \varphi) f(\varphi) g(\theta) \, d\varphi \, d\theta \\ &= \textstyle\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} P_r(\theta - \varphi) f(\varphi) g(\theta) \, d\theta \, d\varphi \\ &= \textstyle\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} P_r(\varphi - \theta) f(\varphi) g(\theta) \, d\theta \, d\varphi, \end{aligned}$$

and thus $\int_{-\pi}^{\pi} (P_r * f)(\theta) g(\theta) \, d\theta = \int_{-\pi}^{\pi} (P_r * g)(\theta) f(\theta) \, d\theta$. Therefore

$$\begin{aligned} \big| \textstyle\int_{-\pi}^{\pi} (P_r * f)(\theta) g(\theta) \, d\theta - \int_{-\pi}^{\pi} f(\theta) g(\theta) \, d\theta \big| &= \big| \textstyle\int_{-\pi}^{\pi} \big[(P_r * g)(\theta) - g(\theta)\big] f(\theta) \, d\theta \big| \\ &\leq 2\pi \|P_r * g - g\|_1 \|f\|_\infty. \end{aligned}$$

By the previous problem the right side tends to 0 as r increases to 1, and the weak-star convergence follows.

21. Let M_f and M_g be upper bounds for $|f|$ and $|g|$ on $[a,b]$. Then

$$\begin{aligned}
&\textstyle\sum_i |f(x_i)g(x_i) - f(x_{i-1})g(x_{i-1})| \\
&\quad \le \textstyle\sum_i |f(x_i)g(x_i) - f(x_i)g(x_{i-1})| + \sum_i |f(x_i)g(x_{i-1}) - f(x_{i-1})g(x_{i-1})| \\
&\quad \le M_f \textstyle\sum_i |g(x_i) - g(x_{i-1})| + M_g \sum_i |f(x_i) - f(x_{i-1})| \\
&\quad \le M_f \|g\|_{BV} + M_g \|f\|_{BV}.
\end{aligned}$$

22. Let us rewrite the given equation $f(x) = f(a) + g_1(x) - g_2(x)$ as $g_2(x) + f(x) - f(a) = g_1(x)$. If $x_i > x_{i-1}$, then subtraction of the values at $x = x_i$ and at $x = x_{i-1}$ gives $g_2(x_i) - g_2(x_{i-1}) + f(x_i) - f(x_{i-1}) = g_1(x_i) - g_1(x_{i-1})$. If $f(x_i) - f(x_{i-1}) \ge 0$, then $f(x_i) - f(x_{i-1}) \le g_1(x_i) - g_1(x_{i-1})$ because g_2 is monotone; if $f(x_i) - f(x_{i-1}) < 0$, then $0 \le g_1(x_i) - g_1(x_{i-1})$ because g_1 is monotone. Therefore $\big(f(x_i) - f(x_{i-1})\big)^+ \le g_1(x_i) - g_1(x_{i-1})$. Summing on i for a partition of $[a,x]$ gives $\sum_{i=1}^n \big(f(x_i) - f(x_{i-1})\big)^+ \le g_1(x) - g_1(a)$. If we take the supremum of the left side and recall that $g_1(a) \ge 0$, we obtain $V^+(f)(x) \le g_1(x) - g_1(a) \le g_1(x)$. Starting similarly from $g_1(x) - f(x) + f(a) = g_2(x)$ and arguing in the same way, we obtain $V^-(f)(x) \le g_2(x) - g_2(a) \le g_2(x)$.

23. Suppose that $V^+(f)$ and $V^-(f)$ are both discontinuous at some x. Then $V^+(f)(x^-) + \epsilon < V^+(f)(x^+)$ and $V^-(f)(x^-) + \epsilon < V^-(f)(x^+)$ for some $\epsilon > 0$. Define

$$g_1(y) = \begin{cases} V^+(f)(y) & \text{for } y < x, \\ V^+(f)(x^-) & \text{for } y = x, \\ V^+(f)(y) - \epsilon & \text{for } y > x, \end{cases}$$

and define $g_2(y)$ similarly except that V^- replaces V^+. Then g_1 and g_2 are both nonnegative, and $g_1 - g_2 = V^+(f) - V^-(f) = f - f(a)$. If g_1 and g_2 are shown to be monotone, Then Problem 22 leads to the contradiction $g_1(y) < V^+(f)(y)$ for $y > x$, and we conclude that $V^+(f)$ and $V^-(f)$ could not have been discontinuous.

In proving monotonicity for g_1, it is necessary to make comparisons only of x with other points y. Let $h > 0$. For points $y > x$, we have $g_1(x+h) = V^+(f)(x+h) - \epsilon \ge V^+(f)(x^+) - \epsilon \ge V^+(f)(x^-) = g_1(x)$. For points $y < x$, we have $g_1(x-h) = V^+(f)(x-h) \le V^+(f)(x^-) = g_1(x)$. Monotonicity for g_2 is proved in the same way.

24. The proof is similar in spirit to the proof of Proposition 6.54.

25. For f, let $y_n = (n+\frac{1}{2})^{-1}\pi^{-1}$, so that $f(y_n)$ is $+(n+\frac{1}{2})^{-1}\pi^{-1}$ if n is even and is $-(n+\frac{1}{2})^{-1}\pi^{-1}$ if n is odd. Compute the sum of the absolute values of the difference of values of f at $y_N, y_{N-1}, \dots, y_1$ and see that this is unbounded as a function of N. The function g has a bounded derivative (even though the derivative is discontinuous), and this is enough to imply bounded variation.

Chapter VII

1. If $g(a_k) = g(b_k)$, then a_k would have to be in E. For the second part an example is $g(x) = x$ on $[0, 1]$; there is only one interval (a_k, b_k), and it is $(0, 1)$.

2. No. Corollary 7.4 applied to I_E shows for almost all x that the quotient $m(E \cap (x-h, x+h))/m((x-h, x+h))$ has to tend to 0 or 1 as h decreases to 0.

3. We may work on a bounded interval I. Let $\epsilon > 0$ be given. If x is in E, then $|h^{-1}(F(x+h) - F(x)| \le \epsilon$ whenever $|h| \le \delta_x$ for some δ_x depending on x. For each such x, fix a positive number r_x with $r_x \le \frac{1}{6}\delta_x$. Associate the set $B(r_x; x)$ to x. Then

$$\mu(B(5r_x; x)) \le \mu((x - 5r_x, x + 5r_x]) = F(x + 5r_x) - F(x - 5r_x) \le 10r_x\epsilon.$$

Applying Wiener's Covering Lemma, we can find disjoint sets $B(r_{x_i}; x_i)$ with $E \subseteq \bigcup_{i=1}^{\infty} B(5r_{x_i}; x_i)$. Then

$$\mu(E) \le \sum_{i=1}^{\infty} \mu(B(5r_{x_i}; x_i)) \le 5\epsilon \sum_{i=1}^{\infty} 2r_{x_i} = 5\epsilon \sum_{i=1}^{\infty} m(B(r_{x_i}; x_i)) \le 5\epsilon m(I).$$

Since I is fixed and ϵ is arbitrary, $\mu(E) = 0$.

4. If F is the function in question, $F - F(0)$ is the distribution function of some Stieltjes measure μ containing no point masses. Proposition 7.8 shows that $\mu(E^c) = 0$ for some countable set E. Since $\mu(\{p\}) = 0$ for each point p, $\mu(E) = 0$ by complete additivity. Thus $\mu = 0$, and F must be constant.

5. For (a), the construction shows that $F'(x) = 0$ for all $x \in C^c$. Then Proposition 7.8 allows us to conclude that μ is singular.

For (b), let F_n be the n^{th} constructed approximation to F (using straight-line interpolations), and let f_n be its derivative (defined except on a finite set and put equal to 0 there). The function f_n is a multiple c_n of the indicator function of the subset C_n of $[0, 1]$ that remains after the first n steps of the construction, and also $m(C_n) = \prod_{k=1}^{n}(1 - r_k)$. Since $F_n(x) = \int_0^x f_n(t)\,dt$ for all x, we have $1 = F_n(1) = c_n \int_0^1 I_{C_n}(t)\,dt = c_n \prod_{k=1}^{n}(1 - r_k)$. Therefore $f_n = \big(\prod_{k=1}^{n}(1 - r_k)\big)^{-1} I_{C_n}$. Put $f = P^{-1} I_C$. The functions f_n converge pointwise to f, and they are uniformly bounded by the constant function P^{-1}. By dominated convergence, $F(x) = \int_0^x f(t)\,dt$ for $0 \le x \le 1$. Therefore F is the distribution function of the measure $f(t)\,dt$.

6. Let E be the second described set. The complement of E has measure 0 by Corollary 7.4. Fix x in E, and let $\epsilon > 0$ be given. Choose a rational r such that $|r - f(x)| < \epsilon$. For $h > 0$,

$$h^{-1} \int_x^{x+h} |f(t) - f(x)|\,dt \le h^{-1} \int_x^{x+h} |f(t) - r|\,dt + h^{-1} \int_x^{x+h} |r - f(x)|\,dt.$$

The second term on the right side equals $|r - f(x)| < \epsilon$, and the first term tends to $|f(x) - r| < \epsilon$ since x is in E. A similar argument applies if $h < 0$.

7. Part (a) is routine, and part (b) follows by adapting part of the argument for Theorem 6.48. In (c), the assumption that x is in the Lebesgue set implies that $\int_{|t|\le h} |f(x-t) - f(x)|\,dt \le hc_x(h)$ for $h > 0$, where $c_x(\cdot)$ is a function that tends to 0 as h decreases to 0. For each of the described pieces of the integral $\int_{|t|\le\pi} K_n(t)|f(x-t)-f(x)|\,dt$, we use one of the two estimates in (a), specifically the estimate $K_N(t) \le N+1$ for the piece with $|t| \le 1/N$ and the estimate $K_N(t) \le c/(Nt^2)$ for all the other pieces. The piece for $1/N$ then contributes $\le (N+1)\int_{|t|\le 1/N} |f(x-t)-f(x)|\,dt \le 2c_x(1/N)$, the piece for $2^{k-1}/N \le |t| \le 2^k/N$ contributes $\le \frac{c}{N}(2^{k-1}/N)^{-2}\int_{2^{k-1}/N\le|t|\le 2^k/N} |f(x-t)-f(x)|\,dt \le \frac{c}{N}(2^{k-1}/N)^{-2}(2^k/N)c_x(2^k/N) = 4\cdot 2^{-k}c_x(2^k/N)$, and finally the piece for $N^{-1/4} \le |t| \le \pi$ contributes $\le \frac{c}{N}N^{1/2}\int_{N^{-1/4}\le|t|\le\pi} |f(x-t)-f(x)|\,dt \le \frac{c}{N}N^{1/2}2\pi(\|f\|_1 + |f(x)|)$. The sum of the estimates is

$$\begin{aligned} &\le 2c_x(1/N) + \sum_{k=1}^{[N^{3/4}]} 4\cdot 2^{-k}c_x(2^k/N) + 2\pi c N^{-1/2}(\|f\|_1 + |f(x)|) \\ &\le 4 \sup_{0<h<N^{-1/4}} c_x(h) + c' N^{-1/2}(\|f\|_1 + |f(x)|), \end{aligned}$$

and this tends to 0 as h decreases to 0. (The use of the shells with 2^{-k} is a device that appears frequently in Zygmund's *Trigonometric Series* and may be regarded as a kind of manual integration by parts.)

8. Since μ is singular, find a Borel set E with $\mu(E) = 0$ and $m(E^c) = 0$. Let $\epsilon > 0$ be given. By regularity of $m + \mu$, choose an open set U containing E such that $(m+\mu)(U - E) < \epsilon$. Then $\mu(U) \le \mu(U-E) + \mu(E) = \mu(U-E) < \epsilon$, and $m(U^c) \le m(E^c) = 0$.

9. About each x in U, there is some $\delta(x)$ such that $(x-h, x+h) \subseteq U$ for $h \le \delta(x)$. Then $\nu((x-h, x+h)) = 0$ for $h \le \delta(x)$, and the limit of this is 0 as h decreases to 0.

11. Since U is open and $\mu_2(U) = 0$, Problem 9 gives

$$\lim_{h\downarrow 0}(2h)^{-1}\mu_2((x-h, x+h)) = 0$$

for all x in U. Since $m(U^c) = 0$, $\lim_{h\downarrow 0}(2h)^{-1}\mu_2((x-h, x+h)) = 0$ for almost every x in $\mathbb{R}^1$. The measure μ_1 has $\mu_1(\mathbb{R}^1) = \mu(U) < \epsilon$, and Problem 10 shows that

$$\begin{aligned} &m\bigl\{x \bigm| \limsup_{h\downarrow 0}\mu_1((x-h, x+h)) > \xi\bigr\} \\ &\qquad\qquad \le m\bigl\{x \bigm| \sup_{h>0}\mu_1((x-h, x+h)) > \xi\bigr\} \le 5\mu_1(\mathbb{R}^1)/\xi < 5\epsilon/\xi. \end{aligned}$$

12. It is enough to handle the case that μ vanishes outside some interval and hence has $\mu(\mathbb{R}^1)$ finite. Combining the estimates for μ_1 and μ_2 gives

$$m\bigl\{x \bigm| \limsup_{h\downarrow 0}\mu((x-h, x+h)) > \xi\bigr\} < 5\epsilon/\xi.$$

Since ϵ is arbitrary, $m\{x \mid \limsup_{h\downarrow 0} \mu((x-h, x+h)) > \xi\} = 0$. Taking the union for $\xi = 1/n$, we conclude that the set where $\limsup_{h\downarrow 0} \mu((x-h, x+h)) > 0$ has measure 0.

To get the better conclusion, the main step is to obtain a bound $10\epsilon/\xi$ for the maximal function formed from the supremum of $\nu((x, x+h))$ or $\nu((x-h, x))$. The proof of Corollary 6.40 shows how to derive this from Problem 10.

Chapter VIII

1. Let $\mathcal{F}$ be the Fourier transform as defined in the text. In each part of the problem, α can be computed by relating matters to the known facts about $\mathcal{F}$, and β can be computed directly from the definitions and Fubini's Theorem.

In (a), we have $\widehat{f}(y) = \int f(x)e^{-ix\cdot y}\,dy = \int f(x)e^{-2\pi i x\cdot(y/(2\pi))}\,dy = \mathcal{F}f(y/(2\pi))$. To obtain $f(x) = \alpha\int \widehat{f}(y)e^{ix\cdot y}\,dy$, we want $f(x) = \alpha\int \mathcal{F}f(y/(2\pi))e^{ix\cdot y}\,dy = (2\pi)^N\alpha\int \mathcal{F}f(y')e^{ix\cdot(2\pi y')}\,dy' = (2\pi)^N\alpha f(x)$. With $f*g(x) = \beta\int f(x-t)g(t)\,dt$, we have $\widehat{f*g}(y) = \beta\iint f(x-t)g(t)e^{-ix\cdot y}\,dt\,dx = \beta\iint f(x-t)g(t)e^{-ix\cdot y}\,dx\,dt = \beta\iint f(x)g(t)e^{-i(x+t)\cdot y}\,dx\,dt = \beta\widehat{f}(y)\widehat{g}(y)$. Thus $\alpha = (2\pi)^{-N}$ and $\beta = 1$.

In (b), we find similarly that $\widehat{f}(y) = (2\pi)^{-N}\mathcal{F}f(y/(2\pi))$, and we are led to $(2\pi)^N(2\pi)^{-N}\alpha = 1$. So $\alpha = 1$. Also, $\beta(2\pi)^N = (2\pi)^{2N}$ and $\beta = (2\pi)^N$.

In (c), we find similarly that $\alpha = (2\pi)^{-N/2}$ and $\beta = (2\pi)^{N/2}$. This normalization has the property that α and β are both 1 if dx is replaced by $dx/(2\pi)^{N/2}$ throughout.

2. This is an operation called "polarization" in linear algebra, and it will be explained further in Chapter XII. Application of the Plancherel formula to $f + cg$, f, and cg gives $\|f + cg\|_2^2 = \|\mathcal{F}(f) + c\mathcal{F}(g)\|_2^2$, $\|f\|_2^2 = \|\mathcal{F}(f)\|_2^2$, and $\|cg\|_2^2 = \|c\mathcal{F}(g)\|_2^2$. We expand the first one in terms of the inner product and subtract the other two to obtain

$$(f, cg)_2 + (cg, f)_2 = (\mathcal{F}(f), c\mathcal{F}(g))_2 + (c\mathcal{F}(g), \mathcal{F}(f))_2.$$

Then $\bar{c}(f, g)_2 + c\overline{(f, g)_2} = \bar{c}(\mathcal{F}(f), \mathcal{F}(g))_2 + c\overline{(\mathcal{F}(f), \mathcal{F}(g))_2}$. Taking $c = 1$ gives $2\,\mathrm{Re}(f, g)_2 = 2\,\mathrm{Re}(\mathcal{F}(f), \mathcal{F}(g))_2$, whereas taking $c = i$ gives $2\,\mathrm{Im}(f, g)_2 = 2\,\mathrm{Im}(\mathcal{F}(f), \mathcal{F}(g))_2$. The result follows.

3. For any f in L^1, we have $Q_\varepsilon * (Q_{\varepsilon'} * f) = P_{\varepsilon+\varepsilon'} * f$ because the Fourier transforms are equal. Also, $(Q_\varepsilon * Q_{\varepsilon'}) * f = Q_\varepsilon * (Q_{\varepsilon'} * f)$ since we have finiteness when the functions are replaced by their absolute values. Moreover, the functions $Q_\varepsilon * Q_{\varepsilon'}$ and $P_{\varepsilon+\varepsilon'}$ are bounded and continuous. Letting f run through an approximate identity formed with respect to dilations and applying Theorem 6.20c, we see that $Q_\varepsilon * Q_{\varepsilon'}(x) = P_{\varepsilon+\varepsilon'}(x)$ for all x.

4. Since P_t is even, $\int_{\mathbb{R}^N}(P_t*f)(x)g(x)\,dx = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} P_t(x-y)f(y)g(x)\,dy\,dx = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} P_t(x-y)f(y)g(x)\,dx\,dy = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} P_t(y-x)f(y)g(x)\,dx\,dy$, and thus

$\int_{\mathbb{R}^N}(P_t * f)(x)g(x)\,dx = \int_{\mathbb{R}^N}(P_t * g)(x)f(x)\,dx$. Therefore

$$\begin{aligned}\Big|\int_{\mathbb{R}^N}(P_t * f)(x)g(x)\,dx - \int_{\mathbb{R}^N} f(x)g(x)\,dx\Big| &= \Big|\int_{\mathbb{R}^N}\big[(P_t * g)(x) - g(x)\big]f(x)\,dx\Big| \\ &\le \|P_t * g - g\|_1\|f\|_\infty.\end{aligned}$$

By Theorem 8.19c the right side tends to 0 as t decreases to 0, and (a) follows.

For (b), part (a) shows for each g with $\|g\|_1 \le 1$ that $\big|\int_{\mathbb{R}^N} f(x)g(x)\,dx\big| = \lim_{t\downarrow 0}\big|\int_{\mathbb{R}^N}(P_t * f)(x)g(x)\,dx\big|$. Since $\big|\int_{\mathbb{R}^N} P_t * f(x)g(x)\,dx\big| \le \|P_t * f\|_\infty\|g\|_1 \le \|P_t * f\|_\infty$, we have

$$\Big|\int_{\mathbb{R}^N} f(x)g(x)\,dx\Big| \le \liminf_{t\downarrow 0}\|P_t * f\|_\infty$$

whenever $\|g\|_1 \le 1$. For any $\epsilon > 0$ with $\|f\|_\infty - \epsilon > 0$, let S_ϵ be the set where $|f|$ is $\ge \|f\|_\infty - \epsilon$. Then $m(S_\epsilon) > 0$. Take E to be any subset of S_ϵ with $0 < m(E) < +\infty$, and let $g(x)$ be $m(E)^{-1}\overline{f(x)}/|f(x)|$ on E and zero elsewhere. This function has $\|g\|_1 \le 1$. Then $\big|\int_{\mathbb{R}^N} fg\,dx\big| = \int_{\mathbb{R}^N} fg\,dx = m(E)^{-1}\int_E |f|\,dx \ge \|f\|_\infty - \epsilon$. Hence $\|f\|_\infty - \epsilon \le \big|\int fg\,dx\big| \le \liminf_{t\downarrow 0}\|P_t * f\|_\infty$. Since ϵ is arbitrary, $\|f\|_\infty \le \liminf_{t\downarrow 0}\|P_t * f\|_\infty$. On the other hand, Theorem 8.19b shows that $\|P_t * f\|_\infty \le \|f\|_\infty$. So we have $\|f\|_\infty \le \liminf_{t\downarrow 0}\|P_t * f\|_\infty \le \limsup_{t\downarrow 0}\|P_t * f\|_\infty \le \|f\|_\infty$. Equality must hold throughout, and (b) is thereby proved.

5. In (a), the set function is a measure by Corollary 5.27. It has $\mu(\mathbb{R}^N)$ equal to $\mu_1(\mathbb{R}^N)\mu_2(\mathbb{R}^N)$ and is therefore a Borel measure. If $\mu_1 = f\,dx$ and $\mu_2 = \mu$, then

$$\begin{aligned}(f * \mu)(E) &= \int_{\mathbb{R}^N}(f\,dx)(E - x)\,d\mu(x) = \int_{\mathbb{R}^N}\int_{E-x} f(y)\,dy\,d\mu(x) \\ &= \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} I_{E-x}(y)f(y)\,dy\,d\mu(x) = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} I_E(x+y)f(y)\,dy\,d\mu(x) \\ &= \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} I_E(y)f(y-x)\,dy\,d\mu(x) = \int_{\mathbb{R}^N}\int_E f(y-x)\,dy\,d\mu(x) \\ &= \int_E\Big[\int_{\mathbb{R}^N} f(y-x)\,d\mu(x)\Big]dy.\end{aligned}$$

In (b), we start with an indicator function and compute that

$$\begin{aligned}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} I_E(x+y)\,d\mu_1(x)\,d\mu_2(y) &= \int_{\mathbb{R}^N}\Big[\int_{\mathbb{R}^N} I_{E-y}(x)\,d\mu_1(x)\Big]d\mu_2(y) \\ &= \int_{\mathbb{R}^N}\mu_1(E-y)\,d\mu_2(y) \\ &= (\mu_1 * \mu_2)(E) = \int_{\mathbb{R}^N} I_E\,d(\mu_1 * \mu_2).\end{aligned}$$

Then we pass to simple functions ≥ 0, use monotone convergence, and finally take linear combinations to get $\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} g(x+y)\,d\mu_1(x)\,d\mu_2(y) = \int_{\mathbb{R}^N} g\,d(\mu_1 * \mu_2)$.

In (c), we actually have $\|P_t * \mu\|_1 = \mu(\mathbb{R}^N)$ for every $t > 0$ by Fubini's Theorem.

Part (d) is handled in the same way as Problem 4a. First one shows that $\int_{\mathbb{R}^N}(P_t * \mu)(x)g(x)\,dx = \int_{\mathbb{R}^N}(P_t * g)(x)\,d\mu(x)$ for g in $C_{\text{com}}(\mathbb{R}^N)$. The resulting estimate is $\big|\int_{\mathbb{R}^N}[(P_t * g)(x) - g(x)]\,d\mu(x)\big| \le \|P_t * g - g\|_{\sup}\,\mu(\mathbb{R}^N)$, and then (a) follows from Theorem 8.19d.

6. Part (a) follows from the same argument as for Proposition 8.1a. In (b), the measure δ with $\delta(\{0\}) = 1$ and $\delta(\mathbb{R}^N - \{0\}) = 0$ has $\widehat{\delta}(y) = 1$ for all y. In (c), we use the result of Problem 5b with $g(x) = e^{-2\pi i x\cdot t}$ and get $\int e^{-2\pi i x\cdot t}\, d(\mu_1 * \mu_2)(x) = \iint e^{-2\pi i(x+y)\cdot t}\, d\mu_1(x)\, d\mu_2(y) = \widehat{\mu_1}(t)\widehat{\mu_2}(t)$. In (d), let $\varphi(x) = P_1(x)$. Then $\widehat{\mu} = 0$ implies $\widehat{\varphi_\varepsilon * \mu} = 0$ for every $\varepsilon > 0$. Since $\varphi_\varepsilon * \mu$ is a function, Corollary 8.5 gives $\varphi_\varepsilon * \mu = 0$ for every $\varepsilon > 0$. By Problem 5d, $\varphi_\varepsilon * \mu$ converges weak-star to μ against $C_{\text{com}}(\mathbb{R}^N)$. Therefore $\int_{\mathbb{R}^N} g\, d\mu = 0$ for every g in $C_{\text{com}}(\mathbb{R}^N)$, and Corollary 6.3 shows that $\mu = 0$.

7. This is the same kind of approximation argument as was done in Corollary 6.17.

8. We calculate that $\sum_{i,j} \widehat{\mu}(x_i - x_j)\xi_i\overline{\xi_j} = \sum_{i,j} \int e^{-2\pi i t\cdot(x_i - x_j)}\xi_i\overline{\xi_j}\, d\mu(t) = \int \big(\sum_{i,j} (e^{-2\pi i t\cdot x_i}\xi_i)\overline{(e^{-2\pi i t\cdot x_j}\xi_j)}\big)\, d\mu(t) = \int \big|\sum_j e^{-2\pi i t\cdot x_j}\xi_j\big|^2\, d\mu(t) \ge 0$.

9. For the set $\{0\}$, the condition is that $F(0)|\xi_1|^2 \ge 0$ for all ξ_1; thus $F(0) \ge 0$. For the set $\{x, 0\}$, the condition is that $F(0)|\xi_1|^2 + F(x)\xi_1\overline{\xi_2} + F(-x)\xi_2\overline{\xi_1} + F(0)|\xi_2|^2 \ge 0$. Taking $\xi_1 = \xi_2 = 1$ shows that $F(x) + F(-x)$ is real; taking $\xi_1 = i$ and $\xi_2 = 1$ shows that $i(F(x) - F(-x))$ is real. Therefore $\overline{F(x)} + \overline{F(-x)} = F(x) + F(-x)$ and $\overline{F(x)} - \overline{F(-x)} = -F(x) + F(-x)$. Adding we obtain $F(-x) = \overline{F(x)}$. Hence $-F(x)\xi_1\overline{\xi_2} - \overline{F(x)}\,\overline{\xi_1}\xi_2 \le F(0)(|\xi_1|^2 + |\xi_2|^2)$. If $F(x) \ne 0$, we put $\xi_1 = -1$ and $\xi_2 = F(x)/|F(x)|$ and obtain $|F(x)| \le F(0)$.

10. $\sum_{i,j} F(x_i - x_j)\Phi(x_i - x_j)\xi_i\overline{\xi_j} = \sum_{i,j} \int F(x_i - x_j)e^{-2\pi i t\cdot(x_i - x_j)}\varphi(t)\xi_i\overline{\xi_j}\, dt = \int \big[\sum_{i,j} F(x_i - x_j)\big(\xi_i e^{-2\pi i t\cdot x_i}\big)\big(\overline{\xi_j} e^{-2\pi i t\cdot x_j}\big)\big]\varphi(t)\, dt \ge 0$.

11. Part (a) follows from the boundedness of F obtained in Problem 9.

In (b), every g in $C_{\text{com}}(\mathbb{R}^N)$ satisfies $0 \le \iint F_0(x - y)\overline{g(x)}g(y)\, dx\, dy = \int (F_0 * g)(x)\overline{g(x)} = \int \widehat{F_0 * g}(y)\overline{\widehat{g}(y)}\, dy = \int \widehat{F}_0(y)\widehat{g}(y)\overline{\widehat{g}(y)}\, dy = \int \widehat{F}_0(y)|\widehat{g}(y)|^2\, dy$.

For (c), if f is in L^2, we can approximate f as closely as we like by a member g of $C_{\text{com}}(\mathbb{R}^N)$. Then $f_0|\widehat{g}|^2 = f_0|\mathcal{F}(f)|^2 + 2f_0\,\mathrm{Re}(\mathcal{F}(f)\overline{(\widehat{g} - \mathcal{F}(f))}) + f_0|\widehat{g} - \mathcal{F}(f)|^2$. We integrate and use the resulting formula to compare $\int f_0|\widehat{g}|^2\, dy$ with $\int f_0|\mathcal{F}(f)|^2\, dy$. By the Schwarz inequality and the Plancherel formula, the absolute value of the difference of these is $\le 2\|f_0\|_{\text{sup}}\|f\|_2\|g - f\|_2 + \|f_0\|_{\text{sup}}\|g - f\|_2^2$. Since $\int f_0|\widehat{g}|^2\, dy$ is ≥ 0, it follows that $\int f_0|\mathcal{F}(f)|^2\, dy \ge 0$ for all f in L^2. Since $\mathcal{F}(f)$ is an arbitrary L^2 function and f_0 is continuous, we conclude that f_0 is ≥ 0.

The integrability in (d) is immediate from Lemma 8.7, and the formula $\int f_0\, dy = F(0)$ follows from the Fourier inversion formula.

12. Let ε_n be a sequence decreasing to 0, let Φ in Problem 11 be the function $e^{-\pi\varepsilon_n^2|x|^2}$, and write F_n for the function $F\Phi$. Then Problem 11d shows that $\mu_n = \widehat{F}_n(y)\, dy$ is a finite Borel measure with $\mu_n(\mathbb{R}^N) = F_n(0) = F(0)$. The Helly–Bray Theorem applies and produces a subsequence of $\{\mu_n\}$ convergent to a finite Borel measure μ weak-star against $C_{\text{com}}(\mathbb{R}^N)$. We shall prove that $F(x) = \int e^{2\pi i x\cdot y}\, d\mu(y)$, i.e., that ν with $\nu(E) = \mu(-E)$ is the desired measure.

For each n, the Fourier inversion formula gives $F_n(x) = \int e^{2\pi i x\cdot y}\widehat{F}_n(y)\, dy = \int e^{2\pi i x\cdot y}\, d\mu_n(y)$. Since $F_n(x) \to F(x)$ pointwise, the result would follow if we could say that the weak-star convergence implies that $\int e^{2\pi i x\cdot y}\,d\mu_n(y) \to \int e^{2\pi i x\cdot y}\,d\mu(y)$. However, $e^{2\pi i x\cdot y}$ is not compactly supported, and an additional argument is needed.

First we extend the weak-star convergence so that it applies to continuous functions vanishing at infinity. If f is such a function, we can find a sequence $\{f_k\}$ in $C_{\text{com}}(\mathbb{R}^N)$ converging to f uniformly. Then

$$\begin{aligned}
&\big|\textstyle\int f\,d\mu_n - \int f\,d\mu\big| \\
&\qquad \le \big|\textstyle\int f\,d\mu_n - \int f_k\,d\mu_n\big| + \big|\int f_k\,d\mu_n - \int f_k\,d\mu\big| + \big|\int f_k\,d\mu - \int f\,d\mu\big| \\
&\qquad \le \|f_k - f\|_{\sup}\,\mu_n(\mathbb{R}^N) + \big|\textstyle\int f_k\,d\mu_n - \int f_k\,d\mu\big| + \|f_k - f\|_{\sup}\,\mu(\mathbb{R}^N).
\end{aligned}$$

Choose k to make $\|f_k - f\|_{\sup}$ small. With k fixed, choose n to make the middle term small. Then the right side is small since the numbers $\mu_n(\mathbb{R}^N)$ are bounded.

This is not quite good enough by itself because $e^{2\pi i x\cdot y}$ does not vanish at infinity. However, averages of it by L^1 functions (i.e., Fourier transforms of L^1 functions) vanish at infinity, and that will be enough for us.

Define $F^{\#}(x) = \int e^{2\pi i x\cdot y}\,d\mu(y)$. We prove that $F^{\#}(x) = F(x)$ for all x. It is enough to prove that $\int F^{\#}\psi\,dx = \int F\psi\,dx$ for all ψ in L^1. Define $\psi^\vee(y) = \int e^{2\pi i x\cdot y}\psi(x)\,dx$. The multiplication formula (for $(\,\cdot\,)^\vee$ instead of $(\,\cdot\,)\widehat{\ }$) and the Riemann–Lebesgue Lemma give

$$\begin{aligned}
\textstyle\int F^{\#}\psi\,dx &= \textstyle\int \psi^\vee\,d\mu(y) = \lim_n \int \psi^\vee\,d\mu_n = \lim_n \int \psi^\vee \widehat{F_n}\,dy \\
&= \textstyle\lim_n \int \psi\,\widehat{F_n}^{\vee}\,dy = \lim_n \int \psi F_n\,dy.
\end{aligned}$$

The right side equals $\int \psi F\,dy$ by dominated convergence since $|F_n(y)| \le |F(y)|$ for all y.

13. Part (a) is easy.

In (b), if χ is a character, then $\sum_x \chi(x) = \sum_x \chi(gx) = \chi(g)\sum_x \chi(x)$. Thus $\sum_x \chi(x) = 0$ if there is some g with $\chi(g) \ne 1$, i.e., if χ is not trivial. If χ and χ' are distinct characters, then $\chi\overline{\chi'}$ is not trivial, and therefore $\sum_x \chi(x)\overline{\chi'(x)} = 0$. The orthogonality implies the linear independence.

In (c), the element 1 of J_m has order m under the group operation of addition. Thus each character χ of J_m must have $\chi(1)$ equal to an m^{th} root of unity. Since 1 generates J_m, $\chi(1)$ determines χ. Thus the listed characters are the only ones.

In (d), any tuple $(n_1, \dots, n_r)$ with $0 \le n_j < m_j$ for $1 \le j \le r$ defines a character by $(k_1, \dots, k_r) \mapsto \prod_{j=1}^r (\zeta_{m_j}^{n_j})^{k_j}$. There are $\prod_{j=1}^r m_j$ distinct characters in this list, and they are linearly independent by (b). Since $\dim L^2(G) = \prod_{j=1}^r m_j$, these characters form a vector-space basis.

14. Since the characters form a basis of $L^2(G)$ as a consequence of Problem 13d, we have $f(t) = \sum_{\chi'} c_{\chi'}\chi'(t)$ for some constants $c_{\chi'}$. Multiply by $\overline{\chi(t)}$ and sum over t to get $\widehat{f}(\chi) = \sum_{\chi'}\sum_t c_{\chi'}\chi'(t)\overline{\chi(t)}$. The orthogonality in Problem 13b shows that this equation simplifies to $\widehat{f}(\chi) = c_\chi \sum_t |\chi(t)|^2 = |G|c_\chi$.

15. $\widehat{f}(\chi) = \sum_{t\in G} f(t)\chi(t) = \sum_{\dot{t}\in G/H} \sum_{h\in H} f(t+h)\dot{\chi}(\dot{t}) = \sum_{\dot{t}\in G/H} F(\dot{t})\dot{\chi}(\dot{t})$ $= \widehat{F}(\dot{\chi})$.

16. The characters of G are the ones with $\chi_n(1) = \zeta_m^n$ for $0 \le n < m$. Such a character is trivial on H if and only if $\chi_n(q) = 1$, i.e., if and only if $\zeta_m^{nq} = 1$; this means that nq is a multiple of m, hence that n is a multiple of p.

The element 1 of H is the element q of G. Thus the question about the identification of the descended characters asks the value of $\chi_n(1)$ when n is a multiple jp of p. The value is $\chi_n(1) = \zeta_m^n = \zeta_{pq}^{jp} = \zeta_q^j$.

If we have computed F on G/H and want to compute $\widehat{F}$ from the definition of Fourier transform, we have to multiply each of the q values of F by the values of each of the q characters of G/H and then add. The number of multiplications is q^2. The actual computation of F from f involves p additions for each of the q values of $\dot{t}$, hence pq additions.

17. $\widehat{f}(\zeta_m^{jp+k}) = \sum_{i=0}^{m-1} f(i)\zeta_m^{(jp+k)i} = \sum_{i=0}^{m-1}(f(i)\zeta_m^{ki})\zeta_m^{jp}$. The variant of f for the number k is then $i \mapsto f(i)\zeta_m^{ki}$. Handling each value of k involves $m = pq$ steps to compute the variant of f and then the $q^2 + pq$ steps of Problem 16. Thus we have $q^2 + 2pq$ steps for each k, which we regard as on the order of $q^2 + pq$. This means $p(q^2 + pq)$ steps when all k's are counted, hence $pq(p+q)$ steps.

Chapter IX

1. Let $r = q/p$, and let r' be the dual index. Regard $|f|^p$ as a product $|f|^p \cdot 1$, and apply Hölder's inequality with $|f|^p$ to be raised to the r power and 1 to be raised to the r' power. Compare with Problem 3 below, which is a more complicated version of the same thing.

2. The inequality is routine if any of the indices is ∞. Otherwise, we have

$$\begin{aligned}\int |fgh|\,d\mu &\le \big(\textstyle\int |fg|^{r'}\,d\mu\big)^{1/r'}\big(\textstyle\int |h|^r\,d\mu\big)^{1/r}\\ &\le \Big(\big(\textstyle\int (|f|^{r'})^{p/r'}\,d\mu\big)^{r'/p}\Big)^{1/r'}\Big(\big(\textstyle\int (|g|^{r'})^{q/r'}\,d\mu\big)^{r'/q}\Big)^{1/r'}\|h\|_r\\ &= \|f\|_p\|g\|_q\|h\|_r.\end{aligned}$$

3. Let us say that $\|f_n\|_p \le C$. Let $\epsilon > 0$ be given. By Egoroff's Theorem, find E with $\mu(E) < \epsilon$ such that f_n tends to f uniformly on E^c. Application of Hölder's inequality with the exponent $r = p/q$ and dual index $r' = p/(p-q)$ to $\int_E |f_n|^q \cdot 1\,d\mu$ gives $\|f_n I_E\|_q \le \big(\int_E |f_n|^{q(p/q)}\,d\mu\big)^{1/p}\big(\int_E 1\,d\mu\big)^{(p-q)/(pq)} \le C\mu(E)^{(p-q)/(pq)} \le C\epsilon^{(p-q)/(pq)}$. Meanwhile, we have

$$\begin{aligned}\|f_n - f\|_q &\le \|f_n - f_n I_{E^c}\|_q + \|f_n I_{E^c} - f I_{E^c}\|_q + \|f I_{E^c} - f\|_q\\ &= \|f_n I_E\|_q + \|(f_n - f)I_{E^c}\|_q + \|f I_E\|_q.\end{aligned}$$

The first term on the right is $\le C\epsilon^{(p-q)/(pq)}$, and so is the third term, by Fatou's Lemma. The middle term tends to 0 as n tends to infinity because of the uniform convergence. Thus $\limsup_n \|f_n - f\|_q \le 2C\epsilon^{(p-q)/(pq)}$. Since ϵ is arbitrary, $\limsup_n \|f_n - f\|_q = 0$.

4. L^1 is 0, and L^∞ consists of the constant functions. All the constant functions give the same linear functional on L^1 because the integral of the product of any constant function and the 0 function is 0.

5. Put $P' = \{f(x) > 0\}$, $N' = \{f(x) < 0\}$, and $Z' = \{f(x) = 0\}$. If E is any measurable subset of Z', then $X = P \cup N$ with $P = P' \cup E$ and $N = N' \cup (Z' - E)$ is a Hahn decomposition. All other Hahn decompositions are obtained by adjusting P and N by taking the symmetric difference of P and of N with any set of μ measure 0.

6. In (a), let X be the positive integers, and let the algebra consist of all finite subsets and their complements; let ν of a finite set be the number of elements in the set, and let ν of the complement of a finite set F be $-\nu(F)$. In (b), use the same X and algebra, define $\nu(\{2k\}) = 2^{-k}$ and $\nu(\{2k-1\}) = -2^{-k}$, and extend ν to be completely additive. In (c), let $X = [0,1]$, let the σ-algebra consist of the Borel sets, and take ν to be Lebesgue measure and μ to be counting measure.

7. Since P_r has L^1 norm 1, the inequality $\|u(r,\,\cdot\,)\|_p \le \|f\|_p$ follows from Minkowski's inequality for integrals. For the limiting behavior as r increases to 1, we extend f periodically and write

$$\begin{aligned} u(r,\theta) - f(\theta) &= \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} P_r(\varphi) f(\theta-\varphi)\,d\varphi - f(\theta) \\ &= \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} P_r(\varphi)[f(\theta-\varphi) - f(\theta)]\,d\varphi, \end{aligned}$$

the second step following since $\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r\,d\varphi = 1$. Applying Minkowski's inequality for integrals, we obtain

$$\|u(r,\,\cdot\,) - f\|_p \le \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} P_r(\varphi)\|f(\theta-\varphi) - f(\theta)\|_{p,\theta}$$

since $P_r \ge 0$. The integration on the right is broken into two sets, $S_1 = (-\delta,\delta)$ and $S_2 = [-\pi,-\delta] \cup [\delta,\pi]$, and the integral is

$$\begin{aligned} &\le \tfrac{1}{2\pi}\textstyle\int_{S_1} P_r(\varphi)\big(\sup_{\varphi\in S_1}\|f(\theta-\varphi) - f(\theta)\|_{p,\theta}\big)\,d\varphi + \tfrac{1}{2\pi}\int_{S_2} P_r(\varphi)2\|f\|_p\,d\varphi \\ &\le \sup_{\varphi\in S_1}\|f(\theta-\varphi) - f(\theta)\|_{p,\theta} + 2\|f\|_p \sup_{\varphi\in S_2} P_r(\varphi). \end{aligned}$$

Let $\epsilon > 0$ be given. If δ is sufficiently small, Proposition 9.11 shows that the first term is $< \epsilon$. With δ fixed, we can then choose r close enough to 1 to make the second term $< \epsilon$.

8. Let p be the dual index to p'. Put $r/R = r'$ in Problem 13 at the end of Chapter IV, so that

$$u(r'R,\theta) = \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} f_R(\varphi)P_{r'}(\theta-\varphi)\,d\varphi$$

for $r' < 1$. Take a sequence of R's increasing to 1, and let $\{R_n\}$ be a subsequence such that $\{f_{R_n}\}$ converges weak-star in $L^{p'}$ relative to L^p. Let the limit be f. For each θ and r', $P_{r'}(\theta - \cdot\,)$ is in L^p, and the equality $u(r'R_n, \theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{R_n}(\varphi)P_{r'}(\theta-\varphi)\,d\varphi$ thus gives $u(r', \theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\varphi)P_{r'}(\theta - \varphi)\,d\varphi$, which is the desired result.

9. If ν is a measure with $0 \le \nu \le \mu$, then $\nu(\{n\}) = 0$ for every n, and hence $\nu(\{\text{integers}\}) = 0$. So $\nu = 0$.

10. Let μ be given on the space X, and consider the set S of all completely additive ν with $0 \le \nu \le \mu$. This contains 0 and hence is nonempty. Order S by saying that $\nu_1 \le \nu_2$ if $\nu_1(E) \le \nu_2(E)$ for all E. If we are given a chain $\{\nu_\alpha\}$, let $C = \sup_\alpha \nu_\alpha(X)$. This is $\le \mu(X)$ and hence is finite. Choose a sequence $\{\nu_{\alpha_k}\}$ from the chain with $\nu_{\alpha_k}(X)$ monotone increasing with limit C.

If $m < n$, let us see that $\nu_{\alpha_m} \le \nu_{\alpha_n}$. Since the ν_α's form a chain, the only way this can fail is to have $\nu_{\alpha_m}(E) > \nu_{\alpha_n}(E)$ for some E and also $\nu_{\alpha_m}(E^c) \ge \nu_{\alpha_n}(E^c)$. But then $\nu_{\alpha_m}(X) > \nu_{\alpha_n}(X)$ by additivity, and this contradicts the fact that $\nu_{\alpha_k}(X)$ is monotone increasing. So $m < n$ implies $\nu_{\alpha_m} \le \nu_{\alpha_n}$.

Define $\nu_0(E) = \lim_k \nu_{\alpha_k}(E)$. Corollary 1.14 shows that ν_0 is completely additive, and certainly $\nu_0 \le \mu$. So ν_0 is an upper bound for the chain. Zorn's Lemma therefore shows that S has a maximal element ν.

Write $\sigma = \mu - \nu$. This is bounded nonnegative additive as a result of the construction. If there were a completely additive λ such that $0 \le \lambda \le \sigma$, then $\nu + \lambda$ would contradict the construction of ν from Zorn's Lemma. Thus σ is purely finitely additive.

11. It is enough to prove that μ is completely additive. If the contrary is the case, then there exists an increasing sequence of sets E_n with union E in the algebra such that the monotone increasing sequence $\{\mu(E_n)\}$ does not have limit $\mu(E)$. Since μ is nonnegative additive, $\mu(E_n) \le \mu(E)$ for all n. Thus $\lim_n \mu(E_n) < \mu(E)$. Since $\nu - \mu$ is nonnegative additive, $\nu - \mu$ similarly has $\lim_n(\nu - \mu)(E_n) \le (\nu - \mu)(E)$. Adding, we obtain $\lim_n \nu(E_n) < \nu(E)$, in contradiction to the complete additivity of ν.

12. Suppose μ is nonnegative bounded additive. Let $\mu = \nu_1 + \rho_1 = \nu_2 + \rho_2$ with ν_1 and ν_2 nonnegative completely additive and with ρ_1 and ρ_2 nonnegative purely finitely additive. Then $\nu_1 - \nu_2 = \rho_2 - \rho_1$. Let $\nu^+ - \nu^-$ be the Jordan decomposition of $\nu_1 - \nu_2$. Since $\nu_1 - \nu_2$ is completely additive, so are ν^+ and ν^-. The equality $\nu^+ - \nu^- = \rho_2 - \rho_1$ and the minimality of the Jordan decomposition together imply that $0 \le \nu^+ \le \rho_2$ and $0 \le \nu^- \le \rho_1$. Problem 11 then shows that $\nu^+ = \nu^- = 0$. Hence $\nu_1 - \nu_2 = 0$, $\nu_1 = \nu_2$, and $\rho_1 = \rho_2$.

13. Let $R = I \times J$ be centered at (x, y). Then $\frac{1}{m(R)}\iint_R |f(u,v)|\,dv\,du = \frac{1}{m(I)}\int_I \Big[\frac{1}{m(J)}\int_J |f(u,v)|\,dv\Big]du \le \frac{1}{m(I)}\int_I f_1(u,y)\,du = f_2(x,y)$. Taking the supremum over R gives $f^{**}(x,y) \le f_2(x,y)$.

14. $\iint |f^{**}(x,y)|^p\,dx\,dy \le \iint |f_2(x,y)|^p\,dx\,dy = \int\Big[\int |f_2(x,y)|^p\,dx\Big]dy \le A_p^p\int\Big[\int |f_1(x,y)|^p\,dx\Big]dy$ by Corollary 9.21. If we interchange integrals and apply

Corollary 9.21 a second time, we see that this is $\leq A_p^{2p} \int \big[\int |f(x,y)|^p \, dy \big] \, dx = A_p^{2p} \|f\|_p^p$.

15. This is done in the style of Corollary 6.39.

16. Let $\Phi_1 \geq 0$ be a decreasing C^1 function on [0, 1] with $\Phi_1'(0) = 0$, $\Phi_1(1) = 1$, and $\Phi_1'(1) = -1$. Define $\Phi_0(x)$ on [0, 1] to be $\Phi_1(x) \big/ (\pi(1+x^2))$ on [0, 1] and to be $1 \big/ (\pi x(1+x^2))$ on $[1, +\infty)$. Then $\Phi(x) = \Phi_0(|x|)$ has the required property.

17. $\sup_{\varepsilon>0} |(\psi_\varepsilon * f)(x)| \leq \sup_{\varepsilon>0}(|\psi_\varepsilon| * |f|)(x) \leq \sup_{\varepsilon>0}(\Phi_\varepsilon * |f|)(x)$, and then $\sup_{\varepsilon>0} |(\psi_\varepsilon * f)(x)| \leq Cf^*(x)$ by Corollary 6.42. Since $\int_{\mathbb{R}^1} \psi(x)\, dx = 0$, the last part of the proof of Corollary 6.42 shows that $\lim_{\varepsilon>0}(\psi_\varepsilon * f)(x) = 0$ a.e. for f in $L^1(\mathbb{R}^1)$. If f is in $L^\infty(\mathbb{R}^1)$ and a bounded interval is specified, we can write f as the sum of an L^1 function carried on that interval and an L^∞ function vanishing on that interval. The L^1 part is handled by the previous case, and the L^∞ part is handled on that bounded interval by Theorem 6.20c.

18. We use the fact that $Q_\varepsilon = h_\varepsilon + \psi_\varepsilon$, where ψ is integrable with integral 0. Since $h_\varepsilon * f$ and $\psi_\varepsilon * f$ are in L^p, so is $Q_\varepsilon * f$. Convolution by an L^1 function such as P_ε is continuous on L^p by Proposition 9.10. With all limits being taken in L^p as $\varepsilon' \downarrow 0$, we have $P_\varepsilon * (Hf) = P_\varepsilon * (\lim(h_{\varepsilon'} * f)) = \lim P_\varepsilon * (h_{\varepsilon'} * f) = \lim P_\varepsilon * (Q_{\varepsilon'} * f - \psi_{\varepsilon'} * f) = \lim P_\varepsilon * (Q_{\varepsilon'} * f) - (\lim P_\varepsilon * \psi_{\varepsilon'}) * f$. The second term on the right side is 0. If we think of P_ε as in L^1 and $Q_{\varepsilon'}$ as in $L^{p'}$, then we have $P_\varepsilon * (Q_{\varepsilon'} * f) = (P_\varepsilon * Q_{\varepsilon'}) * f = Q_{\varepsilon'+\varepsilon} * f = (P_{\varepsilon'} * Q_\varepsilon) * f = P_{\varepsilon'} * (Q_\varepsilon * f)$. Thus $\lim P_\varepsilon * (Q_{\varepsilon'} * f) = \lim P_{\varepsilon'} * (Q_\varepsilon * f) = Q_\varepsilon * f$, and we conclude that $P_\varepsilon * (Hf) = Q_\varepsilon * f$.

19. $\sup_{\varepsilon>0} |(h_\varepsilon * f)(x)| \leq \sup_{\varepsilon>0} |(Q_\varepsilon * f)(x)| + \sup_{\varepsilon>0} |(\psi_\varepsilon * f)(x)| \leq \sup_{\varepsilon>0} |(P_\varepsilon * Hf)(x)| + Cf^*(x) \leq C'(Hf)^*(x) + Cf^*(x)$, the last inequality following from Corollary 6.42 for P_ε. Let $1 < p < \infty$. Then it follows from Corollary 9.21 that $\| \sup_{\varepsilon>0} |h_\varepsilon * f| \|_p \leq C_p(\|Hf\|_p + \|f\|_p)$, and we conclude from Theorem 9.23c that $\| \sup_{\varepsilon>0} |h_\varepsilon * f| \|_p \leq D_p \|f\|_p$. Lemma 9.24 shows that $\lim_{\varepsilon \downarrow 0}(h_\varepsilon * f)(x) = f(x)$ everywhere if f is in a certain dense subspace of L^p, and it follows as in Problem 15 that $\lim_{\varepsilon \downarrow 0}(h_\varepsilon * f)(x) = f(x)$ almost everywhere if f is arbitrary in L^p.

20. Imitating the proof of parts (a) and (b) of Fejér's Theorem (Theorem 6.48), we readily prove that $K_n * f \to f$ in L^p, where K_n is the Fejér kernel. Therefore finite linear combinations of the exponentials are dense in $L^p([-\pi, \pi])$. For each such linear combination f of exponentials, we have $S_n f = f$ for all sufficiently large n, and hence $S_n f \to f$ in L^p for a dense subset of L^p. Using the given estimate on $\|S_n f\|_p$ and the convergence of $S_n f$ on the dense set, we argue as in the proof of Theorem 9.23b to deduce convergence for all f in L^p.

21. Let $F_n(t) = \frac{2\sin(n+\frac{1}{2})t}{t}$ for $0 < |t| \leq \pi$, and extend F_n periodically. Then $\frac{t}{2} F_n(t) = \sin(n+\frac{1}{2})t = (\sin \frac{1}{2}t) D_n(t)$. Since $(t/2) \big/ \sin \frac{1}{2}t = 1 + t\psi(t)$ with $\psi(t)$ bounded above and below by positive constants on $[-\pi, \pi]$, we see that

$D_n(t) - F_n(t) = \left[\frac{\frac{t}{2}}{\sin\frac{1}{2}t} - 1\right]F_n(t) = 2\psi(t)\sin(n+\frac{1}{2})t$. Then the functions $\psi_n(t) = 2\psi(t)\sin(n+\frac{1}{2})t$ have $D_n - F_n = \psi_n$ and $\|\psi_n\|_1$ bounded. By inspection, $F_n - E_n$ equals the function that is $\frac{2\sin(n+\frac{1}{2})t}{t}$ for $|t| < \frac{1}{2n+1}$ and is 0 for $\frac{1}{2n+1} \le |t| \le \pi$. These functions are $\le 2(n+\frac{1}{2})$ for $|t| < \frac{1}{2n+1}$ and are 0 otherwise; so their L^1 norms are bounded. This proves that $D_n - E_n = \varphi_n$ with $\|\varphi_n\|_1 \le C$ for some C.

If $\|T_n f\|_p \le B_p\|f\|_p$, then we have $\|S_n f\|_p = \|D_n * f\|_p = \|E_n * f + \varphi_n * f\|_p \le \|E_n * f\|_p + \|\varphi_n * f\|_p \le B_p\|f\|_p + \|\varphi_n\|_1\|f\|_p$, and we can take $A_p = B_p + C$.

22. We have $2i\sin(n+\frac{1}{2})t = e^{i(n+\frac{1}{2})t} - e^{-(n+\frac{1}{2})t}$. Thus the effect of the operator T_n on f is the sum of two terms $T_n^{(1)}f + T_n^{(2)}f$, one of which is

$$T_n^{(1)}f(x) = \int_{\frac{1}{2n+1}\le|t|\le\pi} \frac{-if(x-t)e^{-i(n+\frac{1}{2})(x-t)}e^{i(n+\frac{1}{2})x}}{t}\,dt.$$

If we regard f as continued periodically to the interval $[-3\pi, 3\pi]$ and we put f equal to 0 outside that interval, then

$$T_n^{(1)}f(x) = e^{i(n+\frac{1}{2})x}((H_\pi - H_{1/(2n+1)})g)(x) \quad \text{for } x \in [-\pi, \pi],$$

where $g(y) = -i\pi f(y)e^{-i(n+\frac{1}{2})(y)}$ on $[-3\pi, 3\pi]$. With A_p as the constant from Theorem 9.23, Theorem 9.23 gives

$$\begin{aligned}\left(\int_{-\pi}^{\pi}|T_n^{(1)}f(x)|^p\,dx\right)^{1/p} &\le \left(\int_{\mathbb{R}}|T_n^{(1)}f(x)|^p\,dx\right)^{1/p}\\ &\le \left(\int_{\mathbb{R}}|H_\pi g|^p\,dx\right)^{1/p} + \left(\int_{\mathbb{R}}|H_{1/(2n+1)}g|^p\,dx\right)^{1/p}\\ &\le 2A_p\left(\int_{\mathbb{R}}|g|^p\,dx\right)^{1/p} \le 2\pi A_p\left(3\int_{-\pi}^{\pi}|f|^p\,dx\right)^{1/p}.\end{aligned}$$

We get a similar estimate for $T_n^{(2)}f$, and the desired estimate for $T_n f$ follows.

23. Define a signed measure ν on $\mathcal{B}$ by $\nu(B) = \int f\,d\mu$. Then ν is absolutely continuous with respect to the restriction of μ to $\mathcal{B}$, and the Radon–Nikodym Theorem yields a function g measurable with respect to $\mathcal{B}$ such that $\nu(B) = \int_B g\,d\mu$ for all B in $\mathcal{B}$. This function g is $E[f|\mathcal{B}]$. Uniqueness is built into the uniqueness aspect of the Radon–Nikodym Theorem.

24. For those n's such that $\mu(X_n) \ne 0$, $E[f|\mathcal{B}]$ may be defined to be equal everywhere on X_n to the constant $\mu(X_n)^{-1}\int_{X_n} f\,d\mu$. For definiteness, $E[f|\mathcal{B}]$ may be defined to be 0 on each X_n with $\mu(X_n) = 0$.

25. The function f satisfies the defining properties (i) and (ii) of $E[f|\mathcal{A}]$.

26. In (a), we identify $E[E[f|\mathcal{B}] \mid \mathcal{C}]$ as $E[f|\mathcal{C}]$. It is measurable with respect to $\mathcal{C}$ and hence satisfies (i) toward being $E[f|\mathcal{C}]$. Any $C \in \mathcal{C}$ has $\int_C E[E[f|\mathcal{B}] \mid \mathcal{C}]\,d\mu = \int_C E[f|\mathcal{B}]\,d\mu$. In turn this equals $\int_C f\,d\mu$ since C is in $\mathcal{B}$. Hence $E[E[f|\mathcal{B}] \mid \mathcal{C}]$ satisfies (ii) toward being $E[f|\mathcal{C}]$.

In (b), we identify $E[f|\mathcal{B}]+E[g|\mathcal{B}]$ as $E[f+g \mid \mathcal{B}]$. It is measurable with respect to $\mathcal{B}$ and hence satisfies (i). For (ii), each B in $\mathcal{B}$ has $\int_B (E[f|\mathcal{B}] + E[g|\mathcal{B}])\,d\mu = \int_B E[f|\mathcal{B}]\,d\mu + \int_B E[g|\mathcal{B}]\,d\mu = \int_B f\,d\mu + \int_B g\,d\mu = \int_B (f+g)\,d\mu$.

In (c), it is enough to handle $f \geq 0$, and then it is enough to handle $g \geq 0$. If $g = I_B$ with $B \in \mathcal{B}$, then we shall identify $I_B E[f|\mathcal{B}]$ as $E[f I_B \mid \mathcal{B}]$. Certainly $I_B E[f|\mathcal{B}]$ satisfies (i). For (ii), each B' in $\mathcal{B}$ has $\int_{B'} I_B E[f|\mathcal{B}]\,d\mu = \int_{B' \cap B} E[f|\mathcal{B}]\,d\mu = \int_{B'\cap B} f\,d\mu = \int_{B'} I_B f\,d\mu$. This handles g equal to an indicator function. Part (b) allows us to handle g equal to a simple function, and monotone convergence allows us to handle g equal to any nonnegative integrable function. (For this last conclusion one needs to use that $f \geq 0$ implies $E[f|\mathcal{B}] \geq 0$, but this is built into the construction via the Radon–Nikodym Theorem.)

In (d), the important thing is that X is a set in $\mathcal{B}$. Then (ii) and (c) successively give $\int_X f E[g|\mathcal{B}]\,d\mu = \int_X E[f E[g|\mathcal{B}] \mid \mathcal{B}]\,d\mu = \int_X E[f|\mathcal{B}]E[g|\mathcal{B}]\,d\mu$. The right side is symmetric in f and g, and hence the left side is also.

Chapter X

1. For (a), the diagonal $\Delta = \{(y, y) \in Y \times Y\}$ is a closed subset of $Y \times Y$ since Y is Hausdorff, and the function $F : X \to Y \times Y$ given by $F(x) = (f(x), g(x))$ is continuous. Therefore $F^{-1}(\Delta)$ is closed.

2. The argument is the same as for Problem 18 in Chapter II.

3. We argue as in the proof of Theorem 2.53. Taking complements, we see that it is enough to prove that the intersection of countably many open dense sets is nonempty. Suppose that U_n is open and dense for $n \geq 1$. Let x_1 be in U_1. Since U_1 is open, local compactness and regularity together allow us to find an open neighborhood B_1 of x_1 with B_1^{cl} compact and $B_1^{\mathrm{cl}} \subseteq U_1$. We construct inductively points x_n and open neighborhoods B_n of them such that $B_n \subseteq U_1 \cap \cdots \cap U_n$ and $B_n^{\mathrm{cl}} \subseteq B_{n-1}$. Suppose B_n with $n \geq 1$ has been constructed. Since U_{n+1} is dense and B_n is nonempty and open, $U_{n+1} \cap B_n$ is not empty. Let x_{n+1} be a point in $U_{n+1} \cap B_n$. Since $U_{n+1} \cap B_n$ is open, we can find an open neighborhood B_{n+1} of x_{n+1} in U_{n+1} such that $B_{n+1}^{\mathrm{cl}} \subseteq U_{n+1} \cap B_n$. Then B_{n+1} has the required properties, and the inductive construction is complete. The sets B_n^{cl} have the finite-intersection property, and they are closed subsets of B_1^{cl}, which is compact. By Proposition 10.11 their intersection is nonempty. Let x be in the intersection. For any integer N, the inequality $n > N$ implies that x_n is in B_{N+1}. Thus x is in $B_{N+1}^{\mathrm{cl}} \subseteq B_N \subseteq U_1 \cap \cdots \cap U_N$. Since N is arbitrary, x is in $\bigcap_{n=1}^{\infty} U_n$.

4. Let Y be a locally compact dense subset of the Hausdorff space X. If y is in Y, let N be a relatively open neighborhood of y such that $N \subseteq K$ with K compact in Y. Since N is relatively open, $N = U \cap Y$ for some open U in X. It will be proved that $N = U$, so that each point of Y has an X open neighborhood, and then Y will be open. The set K is compact in X and must be closed since X is Hausdorff. The points of $U \cap K$ are in Y since $K \subseteq Y$, and hence $U \cap K \subseteq U \cap Y = N$. Consider

a point x of the open set $U - K$. Suppose x is not in Y. Then x is a limit point of Y since Y is dense. Hence the open neighborhood $U - K$ of y contains a point y' of Y. Then y' is in $U \cap Y = N \subseteq K$ and cannot be in $U - K$, contradiction. We conclude that x is in Y. Then x is in $U \cap Y = N$, and $U = N$.

5. First consider any continuous function $f : Y^* \to [0, 1]$ with $f(y_\infty) = 0$. The set of y's with $f(y) > 1/k$ is open and contains y_∞, thus is a compact subset of Y and must be finite. Hence the set of y's with $f(y) = 0$ has a countable complement.

If Z is normal, apply Urysohn's Lemma to A and B, obtaining a continuous $F : Z \to [0, 1]$ with $f(A) = 1$ and $f(B) = 0$. Enumerate the members of X as $x_1, x_2, \ldots$. For fixed n, $f(y) = F(x_n, y)$ is continuous from Y^* to $[0, 1]$ and is 0 at y_∞. Thus $F(x_n, y) > 0$ only on a countable set S_n of y's, and $F(x_n, y) > 0$ for some n at most on the countable set $S = \bigcup_{n=1}^\infty S_n$. If y_0 is not in S, then $x \mapsto F(x, y_0)$ is continuous from X^* to $[0, 1]$, is 0 for every x other than x_∞, and is 1 at x_∞. This contradicts the continuity, and we conclude that Z is not normal.

6. If E is an infinite set with no limit point, then E is closed and each x in E is relatively open. Hence each x has an open set U_x in X with $U_x \cap E = \{x\}$. These open sets and E^c cover X, and there is no finite subcover. Thus X compact implies that each infinite subset has a limit point.

8. Part (a) follows from Problem 7b and Proposition 10.34. For (b), $f^{-1}(-\infty, a)$ is $\varnothing$ if $a < 0$, is $\mathbb{R} - \{0\}$ if $0 \le a < 1$, and is $\mathbb{R}$ if $a \ge 1$; hence it is open in every case. Part (d) follows from (a). For (e), there exists an upper semicontinuous function $\ge f(x)$, namely the constant function everywhere equal to $\sup |f(x)|$. Then (d) shows that the pointwise infimum over all upper semicontinuous functions $\ge f(x)$ meets the conditions on f^-.

9. For (a), we have $Q_f(x) = f^-(x) + (-f)^-(x)$. Both terms on the right are upper semicontinuous, and the sum is upper semicontinuous by Problem 8c. For (c), $f_-(x) \le f(x) \le f^-(x) = Q_f(x) + f_-(x)$. If $Q_f = 0$, then $f_- = f = f^-$ shows that f is continuous with respect to all sets $\{x < b\}$ and all sets $\{x > a\}$. Hence $f^{-1}(a, b)$ is open for every a and b, and f is continuous with respect to the metric topology. Conversely if f is continuous, then the definition makes $f^- = f$ and $(-f)^- = -f$. Therefore $f^- = f_- = f$ and $Q_f = f^- - f_- = 0$.

10. In (a), that subset of pairs is $(A \times A) \cup (B \times B) \cup \{(x, x) \mid x \in X\}$, which is the union of three closed sets and hence is closed. In (b), let X be a Hausdorff space that is not normal, and take A and B to be disjoint closed sets that cannot be separated by open sets.

11. In (a), $q^{-1}q(x) = p_2((\{x\} \times X) \cap R)$, where p_2 is the projection to the second coordinate of $X \times X$. Since $\{x\}$ is closed and X is compact and R is closed, $(\{x\} \times X) \cap R$ is compact. Then $q^{-1}q(x)$ is compact, hence closed, being the continuous image of a compact set.

In (b), we have $p_2((U^c \times X) \cap R) = \{y \in X \mid (x, y) \in R \text{ for some } x \in U^c\} = \{y \in X \mid q^{-1}q(y) \cap U^c \ne \varnothing\} = \{y \in X \mid q^{-1}q(y) \subseteq U\}^c = V^c$. Since U is open, the left side is closed, by the same considerations as in (a). Thus V^c is closed, and V is open.

In (c), let $q(x)$ and $q(y)$ be distinct points of $X/\sim$. By (a), the disjoint subsets $q^{-1}q(x)$ and $q^{-1}q(y)$ are closed. Since X is normal, find disjoint open sets U_1 and U_2 containing $q^{-1}q(x)$ and $q^{-1}q(y)$, respectively. Let $V_1 = \{z \in X \mid q^{-1}q(z) \subseteq U_1\}$ and $V_2 = \{z \in X \mid q^{-1}q(z) \subseteq U_2\}$. These are disjoint sets, and they are open by (b). Then $q(V_1)$ is open in $X/\sim$ because $q^{-1}q(V_1) = V_1$ is open, and similarly $q(V_2)$ is open. The sets $q(V_1)$ and $q(V_2)$ are disjoint because $q^{-1}q(V_1) = V_1$ and $q^{-1}q(V_2) = V_2$ are disjoint. Thus $q(V_1)$ and $q(V_2)$ are the required open sets separating $q(x)$ and $q(y)$.

For (d), part (c) shows that $X/\sim$ is Hausdorff, and therefore its compact subsets are closed. The image of any closed set is X is the image of a compact set, hence is compact and must be closed. For (e), the answer is "no," and part (f) supplies a counterexample. For (f), the function $p : X \to S^1$ is continuous, and Proposition 10.38a produces a continuous function $p_0 : X/\sim \to S^1$ such that $p = p_0 \circ q$, where q is the quotient map. Then p_0 is continuous and one-one from a compact space onto a Hausdorff space and must be a homeomorphism.

12–13. The proofs are the same as in Section II.8.

14. This is proved in the same way as in Problems 13 and 11 in Chapter II.

15. For (a), call the relation $\sim$. This is certainly reflexive and symmetric. For transitivity let $x \sim y$ and $y \sim z$. Then x and y lie in a connected set E, and y and z lie in a connected set F. The sets E and F have y in common, and Problem 13a shows that $E \cup F$ is connected. Thus $x \sim z$. Part (b) is immediate from Problem 13b. For (c), let x be given, and let U be a connected neighborhood of x. Then U lies in the component of x. Thus the component of x is a neighborhood of each of its points and is therefore open.

16. Form the class $\mathcal{C}$ of all functions F as described, including the empty function, and order the class by inclusion; for the purposes of the ordering, each function is to be regarded as a set of ordered pairs. The class $\mathcal{C}$ is nonempty since the empty function is in it. If we have a chain in $\mathcal{C}$, we form the union F of the functions in the chain. We show that F is an upper bound for the chain. To do so, we need to see that the indicated sets cover X. Thus let $x \in X$ be given. Only finitely many sets U in $\mathcal{U}$ contain x, by assumption. Say these are $U_1, \dots, U_n$. If one of these fails to be in the domain of F, then x lies in $\bigcup_{V \in \mathcal{U},\, V \notin \mathrm{domain}(F)} V$, and x is covered. Thus all of $U_1, \dots, U_n$ may be assumed to be in the domain of F. Each U_j is in the domain of some function F_j in the chain, and all of them are in the domain of the largest of the F_j's, say F_0. Since x is not in $\bigcup_{V \in \mathcal{U},\, V \notin \mathrm{domain}(F)} V$, it is not in the larger union $\bigcup_{V \in \mathcal{U},\, V \notin \mathrm{domain}(F_0)} V$. Thus it must be in $\bigcup_{U \in \mathrm{domain} F_0} (U)$. Since $F_0(U)^{\mathrm{cl}} \subseteq U$ for each U, x must lie in some $F_0(U_j)$. Then x lies in $F(U_j)$, and F is an upper bound for the chain.

By Zorn's Lemma let F be a maximal element in $\mathcal{C}$. To complete the argument, we show that every set in $\mathcal{U}$ lies in domain(F). Suppose that U_0 is a set in $\mathcal{U}$ that is not in domain(F). Let U' be the union of all $F(U)$ for U in domain(F) and all V other than U_0 that are not in domain(F). Since F is in $\mathcal{C}$, $U' \cup U_0 = X$. Hence U'^c is a closed subset of the open set U_0. Since X is normal, we can find an open set W such that $U'^c \subseteq W \subseteq W^{\mathrm{cl}} \subseteq U$. If we define $F(U) = W$, then we succeed in enlarging the domain of F, in contradiction to the maximality of F. Hence every member of $\mathcal{U}$ lies in domain(F), as asserted.

17. Form the open sets V_U as in the previous problem. For each U in $\mathcal{U}$, apply Urysohn's Lemma to find a continuous function $g_U : X \to [0, 1]$ with g_U equal to 1 on V_U and equal to 0 on U^c. The open cover $\{V_U\}$ is locally finite since $\mathcal{U}$ is locally finite. Therefore $g = \sum_{U\in\mathcal{U}} g_U$ is a continuous function on X. Since g_U is positive on V_U and the sets V_U cover X, g is everywhere positive. Therefore the functions $f_U = g_U/g$ have the required properties.

18. If $c_0 = 0$, take $F_0 = 0$. If $c_0 \neq 0$, apply Urysohn's Lemma to obtain a continuous function h with values in $[0, 1]$ that is 1 on P_0 and is 0 on N_0, and then put $F_0 = \frac{2}{3}c_0 h - \frac{1}{3}c_0$.

19. On $P_0 \cap C$, g_0 is $\geq c_0/3$ and F_0 is $c_0/3$. Therefore $g_0 - F_0$ is ≥ 0 and $\leq 2c_0/3$. Similarly on $N_0 \cap C$, $g_0 - F_0$ is ≤ 0 and $\geq -2c_0/3$. Elsewhere on C, g_0 and F_0 are both between $-c_0/3$ and $c_0/3$, and hence $|g_0 - F_0| \leq 2c_0/3$. Thus $|g_0 - F_0| \leq 2c_0/3$ everywhere on C. The function F_1 is continuous from X into $\mathbb{R}$, has $|F_1| \leq \frac{2}{3}(\frac{1}{3}c_0)$, and takes a value $c_1 \leq \frac{2}{3}(\frac{1}{3}c_0)$ on $\{x \in C \mid g_1(x) \geq c_1/3\}$ and the value $-c_1$ on $\{x \in C \mid g_1(x) \leq -c_1/3\}$.

20. Iteration produces continuous functions $F_n : X \to \mathbb{R}$ with $|F_n(x)| \leq \frac{1}{3}\big(\frac{2}{3}\big)^n c_0$ for all x in X and $\big|f(x) - \sum_{i=0}^{n-1} F_i(x)\big| \leq \big(\frac{2}{3}\big)^n c_0$ for all x in C. Let $F(x) = \sum_{n=0}^{\infty} F_n(x)$. The series converges uniformly on X by the estimate on $F_n(x)$ and the Weierstrass M test, and Proposition 10.30 shows that F is continuous on X. If we let n tend to infinity in the estimate on $f(x) - \sum_{i=0}^{n-1} F_i(x)$, we see that F and f agree on C. Finally for x in X,

$$|F(x)| \leq \sum_{n=0}^{\infty} |F_n(x)| \leq \sum_{n=0}^{\infty} \tfrac{1}{3}\big(\tfrac{2}{3}\big)^n c_0 = c_0 = \sup_{y\in C} |f(y)|.$$

Thus $|F|$ and $|f|$ have the same supremum.

21. Every open interval is in the base and hence is open. The closed interval $\{a \leq x \leq b\}$ is the complement of the open set $\{x < a\} \cup \{b < x\}$ and is therefore closed.

22. Let $a < b$ be given. If there exists a c with $a < c < b$, then the open sets $\{x < c\}$ and $\{c < x\}$ separate a and b; otherwise the open sets $\{x < b\}$ and $\{a < x\}$ separate them. Hence X is Hausdorff.

Let a and a closed set F be given with a not in F. Since F^c is a neighborhood of a, there exists a basic open set B containing a that is disjoint from F. If B has some element larger than a, let d be such an element; otherwise let d be undefined.

If B has some element smaller than a, let c be such an element; otherwise let c be undefined. If c and d are both defined, then $F \subseteq \{x < c\} \cup \{d < x\}$, while a is in $\{c < x < d\}$. If c is not defined but d is defined, then $F \subseteq \{x < a\} \cup \{d < x\}$, while a is in $B \cap \{x < d\}$. If d is not defined but c is defined, we argue symmetrically. If neither c nor d is defined, then $B = \{a\}$ is open and closed; hence B^c and B are the required open sets separating F and a.

23. Suppose that any nonempty set with an upper bound has a least upper bound, and let E be a set with a lower bound. We are to produce a greatest lower bound. Let F be the set of all lower bounds for E. This is nonempty, and all elements of F are $\leq e$, where e is an element of E. So F has an upper bound. Let c be a least upper bound. We show that c is a greatest lower bound for E.

If c is not a lower bound for E, then E has some e with $e \leq c$, $e \neq c$, i.e., with $e < c$. All f in F have $f \leq e < c$. So e is a smaller upper bound for F, contradiction. Thus c is a lower bound for E. If there is some greater lower bound, say d, then $c < d \leq e$ for all e in E. This implies that d is in F, and hence c is not an upper bound for F.

24. In (a), suppose that Y is nonempty closed and has an upper bound and a lower bound. We are to prove that Y is compact. It is enough to handle a set $Y = [a, b]$. Let an open cover $\mathcal{U}$ of Y be given, and suppose there is no finite subcover. Let E be the set of all x in $[a, b]$ such that some finite subcollection from $\mathcal{U}$ covers $[a, x]$. Then a is in E. Since E is nonempty and has b as an upper bound, the order completeness shows that E has a least upper bound c. Since we are assuming that $\mathcal{U}$ has no finite subcover of $[a, b]$, $E^c \cap [a, b]$ is nonempty. This set has a lower bound, namely a, and therefore it has a greatest lower bound d.

If e is in E and f is in $E^c \cap [a, b]$, then $e \leq f$. So $e \leq d$, and then $c \leq d$. Suppose $c < d$. Then c must be in E. Any x with $c < x < d$ cannot be in E or E^c, and hence there is no such x. Then a finite subclass of $\mathcal{U}$ that covers $[a, c]$, together with a member of $\mathcal{U}$ that contains d, is a finite open subcover for $[a, d]$ and contradicts the fact that d is not in E. Thus $c = d$.

Now suppose that c is in $E^c \cap [a, b]$. Since $c = d$, E has no largest element. Choose a member U of $\mathcal{U}$ containing c, and find a basic open neighborhood B of c contained in U. Then $B \cap E$ must contain some c' with $c' < c$. A finite subclass of $\mathcal{U}$ covers $[a, c']$, and U covers $[c', c]$. Thus c is in E, and we have a contradiction.

We conclude that c is in E. Since $c = d$, $E^c \cap [a, b]$ has no smallest element. Choose a member U of $\mathcal{U}$ containing c, and find a basic open neighborhood B of c contained in U. Then $B \cap (E^c \cap [a, b])$ must contain an element c' with $c < c'$, and then there must be some c'' with $c < c'' < c'$. A finite subclass of $\mathcal{U}$ that covers $[a, c]$, together with the set U, then covers $[a, c'']$ and shows that c'' is in E. This contradicts the fact that c is an upper bound of E.

In (b), let x be given in X. If $a < x < b$ for some a and b, then $[a, b]$ is the required compact neighborhood of x. If x is a lower bound for X and there exists b with $x < b$, then $[x, b]$ is the required compact neighborhood. If x is an upper bound for X and there exists a with $a < x$, then $[a, x]$ is the required compact neighborhood. Since X has at least two members, there are no other possibilities. So X is locally compact.

25. In (a), the sets $\{x < b\}$ and $\{a < x\}$ are open and disjoint, contain a and b respectively, and have union X. Thus X is disconnected.

In (b), suppose that X is order complete and has no gaps. Assume, on the contrary, that U and V are disjoint nonempty open sets with union X. Say that $u < v$ for some u in U and v in V. It will be convenient to assume that u is not the smallest element in X and v is not the largest; when this assumption is not in place, the same line of proof works except that one may below have to use basic open sets of the form $\{r < x\}$ and $\{x < s\}$, as well as $\{r < x < s\}$.

Form the set S of all $x \in X$ with $x \leq v$ and $(x, v] \subseteq V$. This set has u as a lower bound, and we let b be the greatest lower bound. Then $u \leq b \leq v$. First suppose that b is in V. Choose a basic open set $(r, s) \subseteq V$ with $r < b < s$; this is possible by our temporary assumption because V is open. Then $(\max\{u, r\}, v] \subseteq V$. If $\max\{u, r\} < b$, then $\max\{u, r\}$ is in S and b is not a lower bound for S; thus $b \leq \max\{u, r\}$, i.e., $b = u$. This is impossible since b is assumed to be in V. We conclude that b is in U. Choose a basic open set $(r, s) \subseteq U$ with $r < b < s$; again this is possible by our temporary assumption because U is open. Since there are no gaps, we can find s' with $b < s' < s$. Then $\min\{v, s'\}$ is a lower bound for S, and b cannot be the greatest lower bound unless $\min\{v, s'\} \leq b$, i.e., $b = v$. This is impossible since b is assumed to be in U, and we have arrived at a contradiction.

26. As an ordered set, X is the same as $\mathbb{R}$, and hence its order topology is the same as for $\mathbb{R}$, which is connected. In its relative topology, X is disconnected, being the disjoint union of the open sets $[0, 1)$ and $[2, 3)$.

27. The subset $[0, 1)$ is closed, being the intersection of all sets $\{x \mid x \leq y\}$ for $y \in (1, 2]$. Similarly $(1, 2]$ is closed. Hence they are both open, and X is disconnected. It follows immediately from the definition that there are no gaps.

28. If a nonempty subset of points (x, y) is given, let x_0 be the least upper bound of the x's. If no (x_0, y) is in the set, then $(x_0, 0)$ is the least upper bound for the set. If some (x_0, y) is in the set, let y_0 be the least upper bound of the y's. Then (x_0, y_0) is the least upper bound of the set. We conclude that X is order complete. Problem 24a then shows that X is compact. This proves the compactness in (a). There are no gaps, and Problem 25b thus proves the connectedness. For each $x \in [0, 1]$, the set $\{(x, y) \mid 0 < y < 1\}$ is open. Thus we have an uncountable disjoint union of open sets, and X cannot be separable. Part (b) is handled in the same way.

Chapter XI

1. In (a), every compact subset of X is compact when viewed as in X^*, and this gives inclusion in one direction. In the reverse direction it is enough to show that

when U is open in X^*, then $U - \{\infty\}$ is a Borel set in X. Since X is σ-compact, we can choose an increasing sequence of compact sets K_n with $K_n \subseteq K_{n+1}^o$ and $\bigcup_{n=1}^\infty K_n = X$. Then $U \cap K_{n+1}^o$ is open and bounded, hence is a Borel subset of X. The countable union of these sets is U, and hence U is a Borel set. In (b), the Borel sets of X are the countable sets and their complements. However, every subset U of X is open in X and therefore open in X^*. Its complement in X^* is compact and is a Borel set in X^*. Thus U is a Borel set in X^*.

2. Part (a) of the previous problem shows that every open subset of X is a Borel set, and hence every continuous function is a Borel function.

3. Use the regularity to show that the conclusion holds for indicator functions and hence simple functions. Then pass to the limit.

4. Let I_E be an indicator function. Given $\epsilon > 0$, find by regularity a compact set L and an open set U with $L \subseteq E \subseteq U$ and $\mu(U - L) < \epsilon$. The compact set K will be $K = (U - L)^c = L \cap U^c$. Thus consider the restriction of I_E to the compact set K. Let x be in K. If x is in E, then x is in L. The set $U \cap K = L$ is a relatively open neighborhood of x, and I_E is identically 1 on this. Hence the restriction of I_E to K is continuous at the points of E. Similarly if x is in E^c, then x is in U^c. The set $L^c \cap K = U^c$ is a relatively open neighborhood of x, and we argue similarly. This handles indicator functions, and the result for simple functions follows immediately.

Next suppose that f is a real-valued Borel function ≥ 0. Choose an increasing sequence of simple functions $s_n \geq 0$ with limit f. Let $\epsilon > 0$ be given, and find, by Egoroff's Theorem, a Borel set E with $\mu(E^c) < \epsilon$ such that $\lim s_n(x) = f(x)$ uniformly for x in E. Next find, for each n, a compact subset K_n of X with $\mu(K_n^c) < \epsilon/2^n$ such that $s_n\big|_{K_n}$ is continuous. The set $F = E \cap \big(\bigcap_{n=1}^\infty K_n\big)$ has complement of measure $< 2\epsilon$, and the restriction of every s_n to F is continuous. Since $\{s_n\}$ converges to f uniformly on E, the restriction of f to F is continuous. Using regularity once more, we can find a compact subset K_0 of F such that $\mu(F - K_0) < \epsilon$. Then $\mu(K_0^c) < 3\epsilon$, and the restriction of f to K_0 is continuous.

5. In (a), any rotation preserves Euclidean distances and fixes the origin. Since S_{ab} is exactly the set of points whose distance d from the origin has $a < d < b$, S_{ab} is mapped to itself. Part (b) follows from the change-of-variables formula (Theorem 6.32). The determinant that enters the formula is the determinant of the matrix of the rotation and is 1. The first conclusion of (c) is what the change-of-variables formula gives for the transformation to spherical coordinates when applied to the set S_{ab} if we take Fubini's Theorem into account. It yields $\int_{S_{ab}} LF\,dx = \big(\int_a^b r^2\,dr\big)\big(\int_{S^2} Lf\,d\omega\big) = \big(\int_a^b r^2\,dr\big)\big(\int_{S^2} Lf\,d\omega\big)$. Since $\int_a^b r^2\,dr$ is not zero, we can divide by it and obtain the second conclusion of (c). Part (d) is proved by setting it up to be a special case of the uniqueness in Theorem 11.1.

6. In (a), monotonicity of μ gives $\mu(K) \leq \inf_\alpha \mu(K_\alpha)$. Suppose that $<$ holds. Choose by regularity an open set U containing K such that $\mu(U) < \inf_\alpha \mu(K_\alpha)$. The sets K_α^c form an open cover of the compact set U^c, and there is a finite subcover. The

intersection of the complements is one of the sets K_{α_0}, and it has the property that $K_{\alpha_0} \subseteq U$. Monotonicity then gives $\mu(K_{\alpha_0}) \le \mu(U)$, and thus $\inf_\alpha \mu(K_\alpha) \le \mu(U)$, contradiction.

For (b), consider all compact subsets K of X for which $\mu(K) = 1$. The intersection of any two of these is again one by Lemma 11.9. If K_0 is the intersection of all of them, then K_0 is compact, and (a) shows that $\mu(K_0) = 1$. If K_0 contains two distinct points x and y, find disjoint open neighborhoods U_x and U_y. Then $K_0 = (K_0 - U_x) \cup (K_0 - U_y)$ exhibits K_0 as the union of two proper compact sets. At least one of them must have measure 1, and then K_0 is shown *not* to be the intersection of all compact subsets of measure 1.

In (c) let K_0 be any compact G_δ, and choose a decreasing sequence $\{f_n\}$ in $C(X)$ with limit I_{K_0}. Passing to the limit from the formula $\int_X f_n^2\,d\mu = \big(\int_X f_n\,d\mu\big)^2$, we obtain $\mu(K_0) = \mu(K_0)^2$. Thus $\mu(K_0)$ is 0 or 1. By regularity, μ takes only the values 0 and 1, and (b) shows that μ is a point mass.

For (d), apply Theorem 11.1 and obtain the regular Borel measure μ corresponding to ℓ. Then μ has the property in (c) and must be a point mass.

7. The statement for (a) is that $u(r,\theta)$ is the Poisson integral of a signed or complex Borel measure on the circle if and only if $\sup_{0<r<1} \|u(r,\theta)\|_{1,\theta}$ is finite. The necessity is proved in the same way as in Problem 7 at the end of Chapter IX. The sufficiency is proved in the same way as in Problem 8 in that group, except that the weak-star convergence is in $M(\text{circle})$ relative to $C(\text{circle})$. For (b), expand $u(r,\theta)$ in series as in Problem 13 at the end of Chapter IV. Since u is nonnegative, the L^1 norm over any circle centered at the origin is just the integral, and the result of integrating in θ is that the $n = 0$ term is picked out. Thus $\|u(r,\theta)\|_{1,\theta} = c_0$ for every r. The condition in (a) is satisfied, and u is therefore the Poisson integral of a Borel complex measure. Examination of the proof of (a) shows that the complex measure is a measure.

8. Order topologies are always Hausdorff. Since Ω^* has a smallest element and a largest element, Problems 23 and 24 of Chapter X show that Ω^* is compact if every nonempty subset has a least upper bound. Since the ordering for Ω^* has the property that every nonempty subset has a least element, the existence of least upper bounds is satisfied.

9. First we prove that the intersection of any two uncountable relatively closed sets C and D is uncountable. Assume the contrary. Since $C \cap D$ is countable and the countable union of countable sets is countable, there is some countable ordinal ω greater than all members of $C \cap D$. Since C and D are uncountable, we can find a sequence $\omega < \alpha_1 < \beta_1 < \alpha_2 < \beta_2 < \cdots$ such that each α_j is in C and each α_j is in D. The least ordinal γ greater or equal to all members of the sequence is a countable ordinal and has to be a limit point of both C and D. Since C and D are closed, γ is in $C \cap D$. But $C \cap D$ was supposed to have no ordinals greater than ω. This contradiction shows that $C \cap D$ is uncountable, and of course it is relatively closed also.

Now let a sequence of uncountable relatively closed sets C_n be given. By the previous step we may assume that they are decreasing with n. If $\bigcap_{n=1}^{\infty} C_n = C$ is countable, then there is some countable ordinal ω greater than all members of C. Replacing C_n by $C_n \cap \{x \geq \omega\}$ we may assume that the C_n have empty intersection. Let α_n be the least member of C_n. The result is a monotone increasing sequence since the C_n are decreasing. If α is the least ordinal $\geq$ all α_n, then α is a countable ordinal. It is a limit point of each C_n, hence lies in each C_n. The existence of α contradicts the fact that the C_n have been adjusted to have empty intersection. This contradiction shows that $\bigcap_{n=1}^{\infty} C_n$ is uncountable.

10. For additivity the question is whether the union of two sets that fail to meet the condition of the previous problem can meet the condition. The answer is no because the previous problem shows that the intersection of any two sets meeting the condition again meets the condition. The complete additivity is then a consequence of Corollary 5.3 and the result of the previous problem. The measure μ takes on only the values 0 and 1 and yet is not a point mass because one-point sets do not satisfy the defining property for measure 1. Problem 6b therefore allows us to conclude that μ is not regular.

11. Let μ be a Borel measure on X, and let S be the set of all regular Borel measures ν with $\nu \leq \mu$. This contains 0 and hence is nonempty. Order S by saying that $\nu_1 \leq \nu_2$ if $\nu_1(E) \leq \nu_2(E)$ for all E. If we are given a chain $\{\nu_\alpha\}$, let $C = \sup_\alpha \nu_\alpha(X)$. This is $\leq \mu(X)$ and hence is finite. Choose a sequence $\{\nu_{\alpha_k}\}$ from the chain with $\nu_{\alpha_k}(X)$ monotone increasing with limit C. Then $\nu_{\alpha_k}(E)$ is monotone increasing for every Borel set E, and we define $\nu(E)$ to be its limit. The complete additivity of ν follows from Corollary 1.14, and it is easy to check that $\nu_\alpha \leq \nu \leq \mu$ for all α. We have to check that ν is regular. Let $\epsilon > 0$ be given, and choose ν_{α_k} with $\nu_{\alpha_k}(X) \geq \nu(X) - \epsilon$. If E is given, find K and U with $K \subseteq E \subseteq U$, K compact, U open, and $\nu_{\alpha_k}(U - K) < \epsilon$. Then

$$\begin{aligned}\nu_{\alpha_k}(U - K) + \nu((U - K)^c) + \epsilon &\geq \nu_{\alpha_k}(U - K) + \nu_{\alpha_k}((U - K)^c) + \epsilon \\ &= \nu_{\alpha_k}(X) + \epsilon \geq \nu(X) = \nu(U - K) + \nu((U - K)^c),\end{aligned}$$

and hence $\nu_{\alpha_k}(U - K) + \epsilon \geq \nu(U - K)$. Since $\nu_{\alpha_k}(U - K) < \epsilon$, we obtain $\nu(U - K) < 2\epsilon$. Thus ν is regular. The decomposition readily follows.

12. This follows immediately from Proposition 11.20.

13. Let $\mu = \mu_r + \mu_p = \nu_r + \nu_p$ with μ_r and ν_r regular and with μ_p and ν_p purely irregular. Write $\sigma = \mu_r - \nu_r = \nu_p - \mu_p$ in terms of its Jordan decomposition as $\sigma = \sigma^+ - \sigma^-$. Then $\sigma^+ \leq \mu_r$ and $\sigma^- \leq \nu_r$, and hence σ^+ and σ^- are regular by Proposition 11.20. Also, $\sigma^+ \leq \nu_p$ and $\sigma^- \leq \mu_p$, and the definition of "purely irregular" forces σ^+ and σ^- to be 0. Then $\mu_r = \nu_r$ and $\mu_p = \nu_p$.

14. Let μ be as in Problem 10, and suppose that ν is a regular Borel measure with $\nu \leq \mu$. Since $\nu(\{\infty\}) = 0$, Problem 6a shows that $\lim_{\alpha \uparrow \infty} \nu(\{x \geq \alpha\}) = 0$. For each n, let α_n be the least ordinal such that $\nu(\{x \geq \alpha_n\}) \leq 1/n$. The least ordinal $\geq$ all α_n is a countable ordinal β, and $\nu(\{x \geq \beta\}) = 0$. Since $\{x < \beta\}$ is a countable set, $\mu(\{x < \beta\}) = 0$. Therefore $\nu(\{x < \beta\}) = 0$, and we conclude that $\nu = 0$.

16. For the regularity any set in $\mathcal{F}$ is in some $\mathcal{F}_n$. The sets in $\mathcal{F}_n$ are of the form $\widetilde{E} = E \times \big(\bigtimes_{n=N+1}^{\infty} X_n\big)$ with $E \subseteq \Omega^{(n)}$ and $\nu(\widetilde{E}) = \nu_n(E)$. Given $\epsilon > 0$, choose K compact and U open in $\Omega^{(n)}$ with $K \subseteq E \subseteq U$ and $\nu_n(U - K) < \epsilon$. In Ω, $\widetilde{K}$ is compact, $\widetilde{U}$ is open, $\widetilde{K} \subseteq \widetilde{E} \subseteq \widetilde{U}$, and $\nu(\widetilde{U} - \widetilde{K}) < \epsilon$.

17. Let $E = \bigcup_{n=1}^{\infty} E_n$ disjointly in $\mathcal{F}$. Since ν is nonnegative additive, we have $\sum_{n=1}^{\infty} \nu(E_n) \le \nu(E)$. For the reverse inequality let $\epsilon > 0$ be given. Choose K compact and U_n open with $K \subseteq E$, $E_n \subseteq U_n$, $\nu(U_n - E_n) < \epsilon/2^n$, and $\nu(E - K) < \epsilon$. Then $K \subseteq \bigcup_{n=1}^{\infty} U_n$, and the compactness of K forces $K \subseteq \bigcup_{n=1}^{N} U_n$ for some N. Then $\nu(E) \le \nu(K) + \epsilon \le \nu\big(\bigcup_{n=1}^{N} U_n\big) + \epsilon \le \sum_{n=1}^{N} \nu(U_n) + \epsilon \le \sum_{n=1}^{N} \nu(E_n) + 2\epsilon \le \sum_{n=1}^{\infty} \nu(E_n) + 2\epsilon$. Since ϵ is arbitrary, $\nu(E) \le \sum_{n=1}^{\infty} \nu(E_n)$.

18. The key is that Ω is a separable metric space. Every open set is therefore the countable union of basic open sets, which are in the various $\mathcal{F}_n$'s.

Chapter XII

1. In (a), the closed ball is closed and contains the open ball; also every point of the closed ball is a limit point of the open ball since $\|x_1 - x_0\| = r$ implies that $\|[(1-\frac{1}{n})(x_1-x_0)+x_0]-x_0\| = (1-\frac{1}{n})\|x_1-x_0\| < r$ and $\lim_n[(1-\frac{1}{n})(x_1-x_0)+x_0] = x_1$.

For (b), let the closed balls be $B(r_n; x_n)^{\mathrm{cl}}$. If $m \ge n$, then $\|x_m - x_n\| \le r_n$ since $B(r_m; x_m)^{\mathrm{cl}} \subseteq B(r_n; x_n)^{\mathrm{cl}}$. Let $r = \lim_n r_n$. If $r = 0$, then $\{x_n\}$ is Cauchy and hence is convergent. In this case if $x = \lim x_n$, then $\|x - x_n\| \le r_n$ for all n, and hence x is in $B(r_n; x_n)^{\mathrm{cl}}$ for all n. If $r > 0$, fix n_0 large enough so that $r_{n_0} \le 3r/2$. It is enough to show that x_{n_0} is in $B(r_n; x_n)^{\mathrm{cl}}$ for $n \ge n_0$. We may assume that $x_{n_0} \ne x_n$. The members of $B(r_n; x_n)$ are the vectors of the form $x_n + v$ with $\|v\| \le r_n$, and these are assumed to lie in $B(r_{n_0}; x_{n_0})$. Therefore $\|x_n - x_{n_0} + v\| \le r_{n_0}$ for all such v. Take $v = r_{n_0}^{-1} r_n (x_n - x_{n_0})$. Then $r_{n_0} \ge \|x_n - x_{n_0} + v\| = \|(1 + r_{n_0}^{-1} r_n)(x_n - x_{n_0})\| = (1 + r_{n_0}^{-1} r_n)\|x_n - x_{n_0}\|$. Here $r_{n_0}^{-1} r_n \ge (\frac{3}{2} r)^{-1} r = \frac{2}{3}$. So $\|x_n - x_{n_0}\| \le (1 + \frac{2}{3})^{-1} r_{n_0} = \frac{3}{5} r_{n_0} \le \frac{3}{5}\frac{3}{2} r < r \le r_n$, as required.

2. Reduce to the real-valued case, and there use Theorem 1.23 and the remarks at the end of Section A3 of the appendix.

3. Convergence in either case is uniform convergence. For $H^{\infty}(D)$, suppose therefore that $\big\{\sum_{k=0}^{\infty} c_k^{(n)} z_k\big\}$ is a Cauchy sequence in $H^{\infty}(D)$ indexed by n. Write $z = re^{i\theta}$, multiply by $e^{-im\theta}$, and integrate in θ from $-\pi$ to π. The result is that $\big\{c_m^{(n)} r^m\big\}$ is Cauchy in n for each $r < 1$ and each m. Then $\lim_n c_m^{(n)} r^m = c_m r^m$ exists for each r and m. Taking $r = 1/2$, we see that $\lim_n c_m^{(n)} = c_m$ exists for each m. Arguing as in the proof of Theorem 1.37, we see that $f(z) = \sum_{k=0}^{\infty} c_k z_k$ is convergent for $|z| < 1$ and that the sequence of functions $f_n(z) = \sum_{k=0}^{\infty} c_k^{(n)} z_k$ converges to it pointwise. Since $\{f_n\}$ is uniformly Cauchy and pointwise convergent

to f, it converges uniformly to f. For the vector subspace $A(D)$, we have $A(D) = H^\infty(D) \cap C(D^{\mathrm{cl}})$. Hence $A(D)$ is a closed subspace of $H^\infty(D)$.

4. In (a), let us check the triangle inequality. For $y \in Y$, we have $\|a+b+y\| \le \|a+y'\| + \|b+(y-y')\|$ for all $y' \in Y$. Comparing the definition of $\|a+b+Y\|$ with the left side, we obtain $\|a+b+Y\| \le \|a+y'\| + \|b+(y-y')\|$ for all y and y' in Y. Thus $\|a+b+Y\| \le \|a+y'\| + \|b+y''\|$ for all y' and y'' in Y. Taking the infimum over y' and y'' gives the desired conclusion.

In (b), let a Cauchy sequence in X/Y be given. It is enough to prove that some subsequence in convergent. Thus it is enough to prove that if $\{x_n\}$ is a sequence in X with $\|x_n - x_{n+1} + Y\| \le 2^{-n}$, then $\{x_n + Y\}$ is convergent in X/Y. We define a sequence $\{\widetilde{x}_n\}$ in X with $\widetilde{x}_n = x_n - y_n$ and y_n in Y such that $\|\widetilde{x}_n - \widetilde{x}_{n+1}\| \le 2 \cdot 2^{-n}$. It is then easy to check that $\{\widetilde{x}_n\}$ is Cauchy in X and that if x' is its limit, then $\{x_n + Y\}$ tends to $x' + Y$. To define the y_n's, we proceed inductively, starting with $y_1 = 0$. If $y_1, \dots, y_n$ have been defined such that $\|\widetilde{x}_k - \widetilde{x}_{k+1}\| \le 2 \cdot 2^{-k}$ for $k < n$, choose y_{n+1} in Y such that $\|\widetilde{x}_n - x_{n+1} + y_{n+1}\| \le \|x_n - x_{n+1} + Y\| + 2^{-n} \le 2 \cdot 2^{-n}$. Then $\widetilde{x}_{n+1} = x_{n+1} - y_{n+1}$ has $\|\widetilde{x}_n - \widetilde{x}_{n+1}\| \le 2 \cdot 2^{-n}$, and the induction is complete.

5. In (a), we have $c^{\mathrm{tr}} G(v_1, \dots, v_n)\bar{c} = \sum_{i,j} c_i (v_i, v_j)\bar{c}_j = \sum_{i,j} (c_i v_i, c_j v_j) = \big(\sum_i c_i v_i, \sum_j c_j v_j\big) = \big\|\sum_i c_i v_i\big\|^2$. In (b), $G(v_1, \dots, v_n)$ is Hermitian, and thus the finite-dimensional Spectral Theorem says that there exists a unitary matrix $u = [u_{ij}]$ with $u^{-1}G(v_1, \dots, v_n)u$ diagonal, say $= \operatorname{diag}(d_1, \dots, d_n)$. Then $d_j = e_j^{\mathrm{tr}} u^{-1} G(v_1, \dots, v_n) u e_j$, and this, by (a), equals $\big\|\sum_i c_i v_i\big\|^2$ with $\bar{c} = ue_j$. Hence $d_j \ge 0$. In (c), we have $\det G(v_1, \dots, v_n) = \det(u^{-1}G(v_1, \dots, v_n)u) = d_1 d_2 \cdots d_n \ge 0$ with equality if and only if some d_j is 0. If $d_j = 0$, then $\sum_i c_i v_i = 0$ for $\bar{c} = ue_j$, and hence $v_1, \dots, v_n$ is dependent. Conversely if $v_1, \dots, v_n$ is dependent, then $\sum_i c_i v_i = 0$ for some nonzero tuple $(c_1, \dots, c_n)$, and therefore $0 = \big(\sum_i c_i v_i, v_j\big) = \sum_i c_i (v_i, v_j)$ for all j; this equality shows that a nontrivial linear combination of the rows of $G(v_1, \dots, v_n)$ is 0, and hence $\det G(v_1, \dots, v_n) = 0$.

6. A single induction immediately shows the following: $\operatorname{span}\{v_1', \dots, v_k'\} = \operatorname{span}\{u_1, \dots, u_k\}$, v_k' is $\ne 0$, and v_k is defined. Then each v_k has norm 1. If $k < l$, then $(v_l', v_k) = \big(u_l - \sum_{j=1}^{l-1} (u_l, v_j)u_j, v_k\big) = (u_l, v_k) - (u_l, v_k) = 0$. This proves the orthogonality.

7. Define F on each u_α to be the vector v_β given in the statement of the problem, and extend F linearly to a mapping defined on the linear span V of $\{u_\alpha\}$. Corollary 12.8c shows that $\|F(u)\|_{H_2} = \|u\|_{H_1}$ for u in V. Corollary 12.8b shows that V is dense. Proposition 2.47 shows that F extends to a bounded linear operator from H_1 into H_2 satisfying $\|F(u)\|_{H_2} = \|u\|_{H_1}$ for u in H_1. Arguing in the same way with F^{-1} proves that F is onto H_2. The second conclusion follows by using Proposition 12.11.

8. In (a), the boundedness is elementary, and the operator norm is $\|f\|_\infty$. In (b), the adjoint is multiplication by the complex conjugate of f.

9. The linear span V of $\{x_n\}$ is a separable vector subspace. Suppose that it is not dense. Choose by Corollary 12.15 a member $x^* \neq 0$ of X^* with $x^*(V) = 0$. Since $\{x_n^*\}$ is dense, choose a subsequence $\{x_{n_k}^*\}$ with $x_{n_k}^* \to x^*$. Then

$$\|x^* - x_{n_k}^*\| \geq |(x^* - x_{n_k}^*)(x_{n_k})| = |x_{n_k}^*(x_{n_k})| \geq \tfrac{1}{2}\|x_{n_k}^*\|.$$

Since the left side tends to 0, so does the right side. Thus $x_{n_k}^*$ tends to 0, and $x^* = 0$, contradiction.

10. The dual of $C(X)$ is $M(X)$. Define a linear functional x^* on $M(X)$ by $x^*(\rho) = \rho(\{s_0\})$. Then $\|x^*\| = 1$, so that x^* is in $M(S)^*$. Let δ_s denote a point mass at s. If x^* were given by integration with a continuous function f, then we would have $I_{\{s_0\}}(s) = \delta_s(\{s_0\}) = x^*(\delta_s) = \int_S f\,d\delta_s = f(s)$. Thus the only possibility would be $f = I_{\{s_0\}}$, and this is discontinuous.

11. Let X and Y be normed linear with X complete, and let $\{L_n\}$ be a family of bounded linear operators $L_n : X \to Y$ such that $\|L_n(x)\| \leq C_x$ for each x in X. For each y^* in Y^* with $\|y^*\| \leq 1$, the linear functional $y^* \circ L_n$ on X is bounded and has $|y^*(L_n(x))| \leq C_x$. Since X is complete, the Uniform Boundedness Theorem for linear functionals shows that $|y^*(L_n(x))| \leq C\|x\|$ for all x. Taking the supremum over y^* and applying Corollary 12.17, we obtain $\|L_n(x)\| \leq C\|x\|$, as required.

12. For x in X and y in Y, we have

$$\begin{aligned}\|L_n(x) - L_m(x)\| &\leq \|L_n(x-y)\| + \|L_n(y) - L_m(y)\| + \|L_m(y-x)\| \\ &\leq 2C\|x-y\| + \|L_n(y) - L_m(y)\|.\end{aligned}$$

Given $x \in X$ and $\epsilon > 0$, choose y in Y to make the first term $< \epsilon$, and then choose n and m large enough to make the second term $< \epsilon$. It follows that $\{L_n(x)\}$ is Cauchy for each x. Since X' is complete, $L(x) = \lim_n L_n(x)$ exists for all x. Continuity of addition and scalar multiplication implies that L is linear. Then $\|L(x)\| = \lim \|L_n(x)\| \leq \liminf_n \|L_n\|\|x\| \leq C\|x\|$. Hence $\|L\| \leq C$.

13. Proposition 12.1 shows that X^* is a Banach space. We identify the elements x_α in X with their images $\iota(x_\alpha)$ under the canonical map $\iota : X \to X^{**}$. Corollary 12.18 shows that the element $\iota(x_\alpha)$ of X^{**} has $\|\iota(x_\alpha)\| = \|x_\alpha\|$. The hypothesis shows for each x^* that $|(\iota(x_\alpha))(x^*)| = |x^*(x_\alpha)| \leq C_{x^*}$ for a constant C_{x^*} independent of α. Since X^* is complete, the Uniform Boundedness Theorem (Theorem 12.22) shows that $\|\iota(x_\alpha)\| \leq C$ for a constant C independent of α. Applying Corollary 12.18 a second time, we conclude that $\|x_\alpha\| \leq C$ independently of α.

14. For (a), let u and v have $\|u - x\| \leq r$ and $\|v - x\| \leq r$. Then the estimate $\|(1-t)u + tv - x\| = \|(1-t)(u-x) + t(v-x)\| \leq \|(1-t)(u-x)\| + \|t(v-x)\| = (1-t)\|u-x\| + t\|v-x\| \leq (1-t)r + tr = r$ proves the convexity.

For (b), let X be the space of sequences $s = \{s_n\}$ with $\|s\| = \sum_n |s_n|$. Let E_k be the set of sequences with all $s_n \geq 0$, with $\|s\| = 1$, and with $s_j = 0$ for $j \leq k$. If s and t are two sequences with terms ≥ 0, then $\|s + t\| = \|s\| + \|t\|$. The convexity follows, and everything else is easy.

15. Denote open balls in X by B_X and open balls in Y by B_Y. The Interior Mapping Theorem says that $L(B_X(1; 0))$ is open. Hence it contains a ball $B_Y(\epsilon; 0)$. Put $C = \epsilon^{-1}$. By linearity, $L(B_X(Cr; 0)) \supseteq B_Y(r; 0)$ for every $r \geq 0$. Since L is onto Y, we can choose x_0 in X with $L(x_0) = y_0$. Linearity gives $L(B_X(Cr; x_0)) \supseteq B_Y(r; y_0)$. For each y_n, we can take $r = 2\|y_n - y_0\|$ and choose x_n in $B_X(C2\|y_n - y_0\|; x_0)$ with $L(x_n) = y_n$. Since $y_n \to y_0$, $x_n \to x_0$. Also, we have $\|x_n - x_0\| \leq 2C\|y_n - y_0\|$.

In this construction if $y_0 = 0$, we could choose $x_0 = 0$, and then the result follows with $M = 2C$.

If $y_0 \neq 0$, then $\|y_n\| \to \|y_0\| \neq 0$ says that $\|y_n\| \leq \frac{1}{2}\|y_0\|$ only finitely often. For these exceptional n's, we can adjust x_n when $y_n = 0$ so that $x_n = 0$, and then we have $\|x_n\| \leq M\|y_n\|$ for a suitable M and the exceptional n's. For the remaining n's, an inequality $\|x_n\| \leq M\|y_n\|$ is valid as soon as $\{x_n\}$ is bounded, and $\{x_n\}$ has to be bounded since it is convergent.

16. It will be proved that the distance from e to X_0 is ≥ 1. The set X_{00} of all sequences $s_1, s_2 - s_2, s_3 - s_2, \ldots$ such that $\{s_n\}$ is in X is closed under addition and scalar multiplication. Hence it is a dense vector subspace of X_0, and it is enough to prove that $\|e - s\| \geq 1$ for all s in X_{00}. Let s be in X_{00}, and let $c = e - s$. Adding the first n entries gives $c_1 + \cdots + c_n = n - s_n$. Hence $|c_1 + \cdots + c_n| \geq n - \|s\|$. If, by way of contradiction, $\|c\| = 1 - \epsilon$ with $\epsilon > 0$, then $|c_j| \leq 1 - \epsilon$ for all j, and we have $|c_1 + \cdots + c_n| \leq n - n\epsilon$. Thus $n - \|s\| \leq n - n\epsilon$, and we get $n\epsilon \leq \|s\|$, in contradiction to the finiteness of $\|s\|$.

17. This is immediate from Corollary 12.15 and the previous problem.

18. For (a), let $s \geq 0$ have $\|s\| = 1$. Then $\|e - s\| \leq 1$, and so $|x^*(e - s)| \leq 1$. Since $x^*(e) = 1$, this says that $|1 - x^*(s)| \leq 1$. On the other hand, $|x^*(s)| \leq 1$ since $\|s\| \leq 1$. Thus $0 \leq x^*(s) \leq 1$. We can scale this inequality to handle general s.

For (b), the two sequences differ by a member of X_0, on which the Banach limit vanishes identically; then (c) follows by iterated application of (b) since the Banach limit of the 0 sequence is 0.

In (d), let $\epsilon > 0$ be given. By applying (c), we see that we may adjust the sequence so that $\sup_n s_n - \inf_n s_n \leq \epsilon$ and so that the Banach limit is unchanged. By (a), Banach limits preserve order. Since $(\inf s_n)e \leq s \leq (\sup s_n)e$, we have $\inf s_n \leq \mathrm{LIM}_{n\to\infty} s_n \leq \sup s_n$. Since $\sup s_n = (\sup s_n - \limsup s_n) + \limsup s_n \leq (\sup_n - \inf_n) + \limsup_n \leq \limsup_n + \epsilon$, we obtain $\mathrm{LIM}_{n\to\infty} s_n \leq \limsup_n + \epsilon$. Since ϵ is arbitrary, $\mathrm{LIM}_{n\to\infty} s_n \leq \limsup_n$. Similarly $\liminf s_n \leq \mathrm{LIM}_{n\to\infty} s_n$. Conclusion (e) is immediate from (d).

20. The parallelogram law gives

$$2(\|x + y + z\|^2 + \|y - z\|^2) = \|x + 2y\|^2 + \|x + 2z\|^2.$$

If we set $z = 0$ in this identity and then set $y = 0$ in it, we get two relations, one involving an expression for $\|x + 2y\|^2$ and the other involving an expression for $\|x + 2z\|^2$. If we substitute these relations into the displayed equation for the terms $\|x+2y\|^2$ and $\|x+2z\|^2$, we obtain the formula $\|x+y+z\|^2+\|y-z\|^2 = \|x+y\|^2+\|x+z\|^2-\|x\|^2+\|y\|^2+\|z\|^2$. Substitution of $2\|y\|^2+2\|z\|^2-\|y+z\|^2$ for $\|y-z\|^2$ in this formula gives the desired identity.

21. We have

$$\begin{aligned}(x_1+x_2, y) &= \sum_k \tfrac{i^k}{4}\|x_1+x_2+i^k y\|^2 \\ &= \sum_k \tfrac{i^k}{4}(\|x_1+x_2\|^2-\|x_1\|^2-\|x_2\|^2-\|y\|^2) \\ &\quad + \sum_k \tfrac{i^k}{4}\|x_1+i^k y\|^2 + \textstyle\sum_k \tfrac{i^k}{4}\|x_2+i^k y\|^2.\end{aligned}$$

Each term of the first line on the right is 0 because $\sum_k i^k/4 = 0$, and thus the right side simplifies to $(x_1, y)+(x_2, y)$, as required.

22. Induction with the result of the previous problem gives $(nx, y) = n(x, y)$ for every integer $n \geq 0$. Replacing nx by z, we obtain $\frac{1}{n}(z, y) = (\frac{1}{n}z, y)$. Hence $(rx, y) = r(x, y)$ for every rational $r \geq 0$. It follows from the definition of $(\,\cdot\,,\,\cdot\,)$ that $(-x, y) = -(x, y)$ and that if the scalars are complex, $(ix, y) = i(x, y)$. Consequently $(rx, y) = r(x, y)$ if r is in the set D.

23. We are to prove that $|(x, y)| \leq \|x\|\|y\|$, and we may assume that $y \neq 0$. If r is in D, we have

$$0 \leq \|x-ry\|^2 = (x-ry, x-ry) = \|x\|^2 - r(y, x) - \bar{r}(x, y) + |r|^2\|y\|^2.$$

Letting r tend to $(x, y)\big/\|y\|^2$ through members of D, we obtain

$$0 \leq \|x\|^2 - 2|(x, y)|^2\big/\|y\|^2 + |(x, y)|^2\|y\|^2\big/\|y\|^4 = \|x\|^2 - |(x, y)|^2\big/\|y\|^2,$$

and it follows that $|(x, y)| \leq \|x\|\|y\|$.

24. The Schwarz inequality gives

$$|r(x, y) - (cx, y)| = |(rx - cx, y)| \leq \|(r-c)x\|\|y\| = |r-c|\|x\|\|y\|.$$

As r tends to c through D, the right side tends to 0, and the left side tends to $|c(x, y) - (cx, y)|$. Hence $c(x, y) = (cx, y)$.

SELECTED REFERENCES

Adams, M., and V. Guillemin, *Measure Theory and Probability*, Wadsworth & Brooks/Cole Advanced Books and Software, Monterey, CA 1986; corrected reprint of the original, Birkhäuser Boston, Inc., Boston, 1996.

Bartle, R. G., *The Elements of Real Analysis*, John Wiley & Sons, New York, 1964; second edition, 1976.

Bartle, R. G., and D. R. Sherbert, *Introduction to Real Analysis*, John Wiley & Sons, Inc., New York, 1982; second edition, 1992.

Coddington, E. A., and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw–Hill Book Company, New York, 1955.

Doob, J. L., *Measure Theory*, Springer-Verlag, New York, 1994.

Duistermaat, J. J., and J. A. C. Kolk, *Multidimensional Real Analysis*, Vol. I, Cambridge University Press, Cambridge, 2004.

Duistermaat, J. J., and J. A. C. Kolk, *Multidimensional Real Analysis*, Vol. II, Cambridge University Press, Cambridge, 2004.

Dunford, N., and J. T. Schwartz, *Linear Operators*, Part I, Interscience Publishers, Inc., New York, 1958.

Halmos, P. R., *Measure Theory*, D. Van Nostrand Company, Inc., Princeton, 1950; reprinted, Springer-Verlag, New York, 1974.

Halmos, P. R., *Naive Set Theory*, D. Van Nostrand Company, Inc., Princeton, 1960; reprinted, Springer-Verlag, New York, 1974.

Hayden, S., and J. F. Kennison, *Zermelo–Fraenkel Set Theory*, Charles E. Merrill Publishing Company, Columbus, 1968.

Ince, E. L., *Ordinary Differential Equations*, Longmans, Green and Co., London, 1927; reprinted Dover Publications, New York, 1944.

Jackson, D., *Fourier Series and Orthogonal Polynomials*, Carus Mathematical Monographs, vol. 6, Mathematical Association of America, no city listed, 1941.

Kaplan, W., *Ordinary Differential Equations*, Addison–Wesley Publishing Co., Reading, MA, 1958.

Kelley, J. L., *General Topology*, D. Van Nostrand Company, Inc., Princeton, 1955; reprinted, Springer-Verlag, New York, 1975.

Lebesgue, H., *Leçons sur l'Integration et la Recherche des Fonctions Primitives*, Gauthiers-Villars, Paris, 1904; second edition, 1928; second edition reprinted by Éditions Jacques Gabay, Sceaux, 1989.

Munroe, M. E., *Introduction to Measure and Integration*, Addison–Wesley Publishing Company, Inc., Cambridge, MA, 1953; second edition, *Measure and Integration*, Addison–Wesley Publishing Company, Reading, MA, 1971.

Riesz, F., and Béla Sz.-Nagy, *Functional Analysis*, Frederick Ungar Publishing Co., New York, 1955.

Royden, H. L., *Real Analysis*, The Macmillan Company, New York, 1963; second edition, Macmillan Publishing Co., 1968; third edition, 1988.

Rudin, W., *Principles of Mathematical Analysis*, McGraw–Hill Book Company, New York, 1953; second edition, 1964; third edition 1976.

Rudin, W., *Real & Complex Analysis*, McGraw–Hill Book Company, New York, 1966; second edition 1974; third edition 1987.

Saks, S., *Theory of the Integral*, second revised edition, G. E. Stechart and Company, New York, 1937; reprinted, Hafner Publishing Company, New York, no date, and Dover Publications, New York, 1964.

Spivak, M., *Calculus on Manifolds*, Benjamin/Cummings Publishing Company, Menlo Park, CA, 1965.

Stein, E. M., *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.

Stein, E. M., and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, Princeton, 1971.

Stone, M. H., "A generalized Weierstrass approximation theorem," *Studies in Modern Analysis*, R. C. Buck (ed.), Mathematical Association of America, distributed by Prentice–Hall, Inc., Englewood Cliffs, NJ, 1962.

Williamson, R. E., and H. F. Trotter, *Multivariable Mathematics*, Prentice–Hall, Inc., Englewood Cliffs, NJ, 1974; second edition 1979; third edition 1996.

Zygmund, A., *Trigonometric Series*, Vol. I, second edition, Cambridge University Press, Cambridge, 1959.

Zygmund, A., *Trigonometric Series*, Vol. II, second edition, Cambridge University Press, Cambridge, 1959.

INDEX OF NOTATION

See also the list of Standard Notation on page xxi. In the list below, items are alphabetized according to their key symbols. For letters the order is italic lower case, Roman lower case, italic upper case, Roman upper case, script, and blackboard bold. Next come items whose key symbol is Greek, and then come items whose key symbol is a nonletter. The last of these are grouped by type.

a_n, 61
arccos, 79
arcsin, 51
arctan, 51
$A(D)$, 523
b_n, 61
$B(r; x)$, 83
$B(S)$, 87, 281, 521
$B(S, \mathbb{C})$, 87
$B(S, \mathbb{R})$, 87
$\mathcal{B}_1$, 297
$\mathcal{B}_N$, 297
$\mathcal{B}(X)$, 488
$\mathcal{B}(X, Y)$, 284, 524
$\mathcal{B}_0(X)$, 504
$\mathcal{B}_N(K)$, 318
$\mathcal{B}_N(V)$, 315
c, 522
c_0, 522
c_n, 61, 335
card, 577
A^{cl}, 93
cos, 47
C^∞, 142
C^k, 142
$C(S)$, 100, 281, 464, 521
$C(S, \mathbb{C})$, 100, 512
$C(S, \mathbb{R})$, 100, 512
$C_0(X)$, 509, 521
$C^\infty(E)$, 142
$C^k(E)$, 142
$C_{\mathrm{com}}(X)$, 300, 314, 486, 521
$C^N([a, b])$, 522
d, 41
d_2, 75
$d(x, y)$, 83
dx, 298
$d\omega$, 326, 491, 517
D, 384
D_j, 385
D_N, 69
e, 47
e^A, 148
e_j, 566
exp, 47, 148
E_x, 267
E^y, 267
$E[f|\mathcal{B}]$, 440
F_σ, 492
$\mathcal{F}$, 374
$\mathbb{F}$, 135
$\mathbb{F}^n$, 565
G_δ, 492
H, 397
H_1, 401
$H^\infty(D)$, 523

inf, 3, 6
I_E, 171, 241
$\mathcal{I}_E(s)$, 242
$J_p(t)$, 225
K_N, 71
$\mathcal{K}(X)$, 490
ℓ^2, 87
ℓ_n^p, ℓ^p, 522
lim, 5, 98
lim inf, 7
lim sup, 7
log, 49
L^1, 89, 281
L^2, 89, 281
L^∞, 281
L_n, 544
$L(y)$, 208
$L(P, f)$, 27, 162
$L^p(X)$, $L^p(X, \mathcal{A}, \mu)$, $L^p(X, \mu)$, 281, 522
m, 236, 298
$m_R(f)$, 162
$M(X)$, 522
$M(X, \mathbb{C})$, 514
$M(X, \mathbb{R})$, 512
$M_R(f)$, 162
A^o, 92
$\mathrm{osc}_g(x_0)$, 119, 167
p', 411
$X = P \cup N$, 419
$P(\lambda)$, 208
$P(x, t)$, $P_t(x)$, 395
$P_r(\theta)$, 392
$Q(D)$, 384
$Q(x, \varepsilon)$, $Q_\varepsilon(x)$, 398
$Q_n(a_1, \ldots, a_N)$, 302
$\mathbb{Q}^\times$, 134
$\mathbb{Q}_p$, 133
$\mathbb{Q}_p^\times$, 134
$\mathcal{R}[a, b]$, 27
$\mathcal{R}(A)$, 162
$\mathcal{R}\int_A f\,dx$, 318
$\mathbb{R}^*$, 6, 85
$s_N(f; x)$, 62, 335
sin, 47
sup, 2, 5
S^n, 126, 326
$S(P, \{t_i\}, f)$, 39
$S(P, \{t_R\}, f)$, 163
$\mathcal{S}$, 385
$\mathcal{S}(\mathbb{R}^N)$, 385
Tr X, 149
$\mathcal{T}$, 442
$\mathcal{T}^*$, 455
u_j, 566
$U(P, f)$, 27, 162
$V(f)$, 355
$V^*(f)$, $V^-(f)$, 347
$\mathbb{Z}_p$, 134

Greek

$\Gamma(s)$, 323
Δ, 181, 392
Δx_i, 27
$E \Delta F$, 233
$\zeta(s)$, 390
ι, 541
λ_f, 350
μ^*, 254, 255
μ_*, 254, 255
μ_f, 350
$\mu(P)$, 27, 161
$[d\mu]$, 248
$\nu = f\,d\mu$, 252
σ_n, 53
σ_r, 54
τ_x, 303
φ_ε, 312
Ω, 292
Ω_N, 326

Isolated symbols and signs
$\in$, 554, 555
∞, 5
$\subseteq$, $\supseteq$, 555
$+\infty$, $-\infty$, 6
r^+, r^-, 347
f^+, f^-, 239
ν^+, ν^-, 418
$\sim$, 62, 335
$\ll$, 420
$\oplus$, 528

Subscripts and superscripts
X^*, 128, 444, 455, 524
$f^*(x)$, 328
$M^\perp$, 527
L^*, 536
x^*, 537
2^A, 555

Operations on sets
$\mathcal{A} \times \mathcal{B}$, 266
$X/\sim$, 445, 471
$\uparrow$, $\downarrow$, 268
B^A, 556
$\bigcup_{x\in S} A_x$, $\bigcap_{x\in S} A_x$, $\bigtimes_{x\in S} A_x$, 557

Operations on functions
f^{-1}, 558
$f * g$, 180, 306, 334
$\mu \times \nu$, 270
$\widehat{f}$, 374
$\overline{\int_a}^b f\,dx$, $\underline{\int}_a^b f\,dx$, 27
$\overline{\int}_A f\,dx$, $\underline{\int}_A f\,dx$, 162
$\int_a^b f\,dx$, 27, 30
$\int_A f\,dx$, 162
$\int_E f\,d\mu$, 242
$\int_E f(x), d\mu(x)$, 242

Specific functions
$(\cdot\,,\,\cdot)_2$, 75
$\|\cdot\|_2$, 75, 279
$|\cdot|$, 84, 136, 138, 161, 239, 514
$x \cdot y$, 84
$(\cdot\,,\,\cdot)$, 89, 523
$\|\cdot\|$, 90, 136, 279, 514, 521, 523
$|\cdot|_p$, 133
$\|\cdot\|_1$, 279
$\|\cdot\|_\infty$, 279
$\|\cdot\|_{\sup}$, 281
$\|\cdot\|_{BV}$, 347
$\|\cdot\|_p$, 410

Intervals
(a, b), 3
$[a, b]$, 3
$[a, b)$, 3
$(a, b]$, 3

Other
$[M_{ij}]$, 566
$(X, \mathcal{T})$, 442
$a_n \to a$, 5, 98
$\frac{\partial f_i}{\partial x_j}$, 139
$\sum_{j=1}^n a_j(x_1, \ldots, x_n)\frac{\partial}{\partial x_j}$, 199
$(X, \mathcal{A}, \mu)$, 238
$(D, \leq)$, 465
$\{x_\alpha\}$, 465
$\alpha \mapsto x_\alpha$, 465

INDEX

a.e., 248
Abel summable, 54
Abel sums, 54
Abel's Theorem, 54
Abelian theorem, 55
absolute value, 563
absolutely continuous
  measure, 420
  monotone function, 369
  Stieltjes measure, 369
abstract rectangle, 266, 297
additive set function, 234
  bounded, 418
adjoint, 536
Alaoglu's Theorem preliminary form, 284, 509
algebra, 124
algebra of sets, 232
almost every, 248
almost everywhere, 248
almost period, 131
almost periodic, 131
  Bochner, 131
  Bohr, 131
alternating series test, 19
approximate identity, 58, 60, 312
arccosine, 79
archimedean property, 4
arcsine, 51
arctangent, 51
area of sphere, 326
arithmetic operations, 104
arithmetic-geometric mean inequality, 182
Ascoli's theorem, 22, 121, 229
Ascoli–Arzelà Theorem, 478
associated Stieltjes measure, 344
Axiom of Choice, 557
axiom of countability, 450

Baire Category Theorem, 118, 120, 481
Baire function, 505
Baire measurable function, 505
Baire measure, 506
Baire set, 504
ball, open, 83
Banach limit, 550
Banach space, 521
  quotient, 549
  reflexive, 541
Banach–Steinhaus Theorem, 543
base
  countable local, 450
  for metric space, 106
  for topology, 446
  local, 450
basis
  of vector space, 573
  standard, 566
Bessel function, 225
Bessel's equation, 224
Bessel's inequality, 67, 335, 531
binomial coefficient, xxi
binomial series, 52, 56
Bochner almost periodic, 131
Bochner's Theorem, 406
Bohr almost periodic, 131
Bolzano–Weierstrass property, 108, 109
Bolzano–Weierstrass Theorem, 9, 108
Borel complex measure, regular, 514
Borel function, 297, 316, 318, 504
Borel measurable, 239, 297
Borel measurable function, 504
Borel measure 299, 316, 318, 490
  purely irregular, 519
  regular, 299, 316, 318, 490
Borel set, 237, 297, 315, 318, 488, 504
Borel signed measure, regular, 512
bound
  greatest lower, 3

least upper, 2
lower, 3
upper, 2
bounded, 478
bounded additive set function, 418
bounded, essentially, 280
bounded function, 87
bounded interval, 3
bounded linear operator, 283, 523
bounded operator, 283
bounded, pointwise, 22, 121, 478
bounded set, 3, 490
bounded, uniformly, 22, 121, 478
bounded variation, 346, 355

C^k function, 142
C^∞ function, 142
canonical expansion of simple function, 241
canonical extension of measure, 261
canonical map, 541
Cantor 579
Cantor diagonal process, 24, 461
Cantor function, 344
Cantor measure, 354, 519
Cantor set, 118, 290, 343, 519
standard, 118
cardinal number, 577
cardinality, 577
Cartesian product, 556, 557
Cauchy criterion, 10
uniform, 18
Cauchy Integral Theorem, 378
Cauchy sequence, 9, 112
uniformly, 18
Cauchy–Peano Existence Theorem, 229
Cauchy–Riemann equations, 379, 398
Cauchy–Schwarz inequality, 563
Cesàro summable, 53
Cesàro sums, 53, 371
chain, 573
chain rule, 144
change-of-variables formula, 37, 172, 320
character, multiplicative, 407
characteristic function, 171, 241
characteristic polynomial, 208, 214
Chebyshev's inequality, 351
class, 555
equivalence, 564
class C^k, 142
class C^∞, 142
closed geometric rectangle, 161
Closed Graph Theorem, 548
closed interval, 3, 6
closed map 482
closed rectangle, 161
closed set, 3, 92, 442
closed subspace, 102
closed unit disk, 126
closed vector subspace, 104
closure, 93, 442
cofactors, 568
cofinal, 468
collection, 555
compact metric space, 108
compact set, 108, 453
compact topological space, 453
compactification, one-point, 456
complement, 555
complete metric space, 113
completely additive set function, 234
σ-finite, 237
completeness of L^p, 285, 414
completion of measure space, 262, 291
completion of metric space, 128, 133
complex conjugation, 563
complex Euclidean space, 85
complex measure, regular Borel, 514
component, 483
component rectangles, 161
components of a function, 139
composition, 558
conditional expectation, 439
conjugate Poisson integral, 398
conjugate Poisson kernel, 398
conjugation, complex, 563
connected, 482
locally, 131
locally pathwise, 131
connected component, 483
connected metric space, 115
connected set, 115
constant coefficients, 208, 211
content 0, 167
continuity at a point, 95
continuity, uniform, 111
continuous derivative on a closed interval, 561
continuous from the left, 339
continuous from the right, 339

continuous function, 10, 96, 443
  uniformly, 11
continuous linear functional on L^p, 425
continuous periodic, 67
continuous singular Stieltjes measure, 368
contraction mapping, 131
contraction mapping principle, 131
converge, 463, 466
convergence
  norm, 284
  uniform, 17
  weak-star, 284, 355, 405, 437
convergent sequence, 5, 97
convolution, 58, 180, 306, 334, 417
  of measures, 405
cosine, 47
countable, xxi, 106
countable local base, 450
countable ordinals, 292
countably additive set function, 234
counting measure, 236
cover, 106, 151
Cramer's rule, 568
critical point, 323
critical value, 323
curve, 199
  integral, 199
cut, 2

defined implicitly, 153
definite, 89
degree of polynomial, 568
delta mass, 342
delta measure, 342
dense, 106
dense set, 450
derivative, 140
  at endpoint of interval, 561
  on a closed interval, 561
  partial, 139
determinant, 567
  Wronskian, 205
diadic cube, 302
diagonal process, 24
diffeomorphism, 160
difference, 555
differentiable, 139
  infinitely, 46
  periodic, 67
differential equation
  existence theorem for ordinary system, 189, 229
  first-order ordinary linear, 228
  ordinary, 183
  ordinary homogeneous linear, 184, 201
  ordinary inhomogeneous linear, 201
  ordinary linear, 184, 201
  system of ordinary, 188
  uniqueness theorem for ordinary system, 193
differential, total, 153
differentiation, implicit, 153
differentiation of integrals, 35, 329, 363, 370
  strong, 431, 438
differentiation of monotone functions, 359
differentiation of series of monotone functions, 362
dilation, 303
Dini's test, 70, 336
Dini's Theorem, 78, 132, 480, 492
direct image, 559
directed set, 465
Dirichlet kernel, 69
Dirichlet–Jordan Theorem, 348
discrete metric, 87
discrete topology, 443
disk, 126, 181
distance, 41, 75, 83
  to a set, 96
distribution function, 339, 351
divide, 568
Division Algorithm, 569
divisor, 568
  greatest common, 570
domain, 556
Dominated Convergence Theorem, 253
dot product, 84, 566
doubly infinite sequence, 5
dual group of finite abelian group, 407
dual index, 411, 525
dual of normed linear space, weak-star topology, 444
dual space, 284, 524

Egoroff's Theorem, 291, 292, 517
element, 554, 555
elementary matrix, 174
elementary set, 233
entity, 554

equicontinuous, 22, 121, 227, 230, 477, 478
uniformly, 22, 121
equivalence class, 564
equivalence relation, 564
essential bound, 280
essential supremum, 280
essentially bounded, 280
Euclidean algorithm, 570
Euclidean norm, 84, 85
Euclidean space, 84, 521
Euler's equation, 222
eventually, 97, 463, 465
everywhere dense, 106
existence theorem
integral curves, 199
system of ordinary differential equations, 189, 229
expansion by cofactors, 568
expectation, conditional, 439
exponential, 47
of a matrix, 148
extended real number, 6, 85
Extension Theorem, 238, 253, 271, 489, 502

F_σ, 492
factor of polynomial, 568
Factor Theorem, 569
family, 555
fast Fourier transform, 391, 407
Fatou's Lemma, 252
Fatou's Theorem, 334, 395
Fejér kernel, 71
Fejér's Theorem, 72, 337, 371
field of scalars, 279
filter, 294
finest topology, 445
finite abelian group, Fourier analysis on, 407
finite interval, 3
finite limit for derivative at endpoint, 561
finite measure space, 238
finite-intersection property, 109, 454
first axiom of countability, 450
first category, 119
first countable, 450
first-order linear ordinary differential equation, 228
flip, 174
Fourier coefficient, 62, 335
Fourier inversion formula, 380
Fourier series
almost everywhere convergence of Cesàro sums, 371
Bessel's inequality, 67, 335
convergence in L^p, 439
convolution, 180
Dini's test, 70, 336
Dirichlet–Jordan Theorem, 348
divergence for a continuous function, 544
failure of some sequence vanishing at infinity to occur, 547
Fejér's Theorem, 72, 337
localization, 71
Parseval's Theorem, 74, 338
Riemann–Lebesgue Lemma, 67, 335
Riesz–Fischer Theorem, 338
uniqueness theorem, 73, 77, 338
with harmonic functions, 227
with Lebesgue integral, 335
with Riemann integral, 62
Fourier transform, 374
fast, 391, 407
for finite abelian group, 407
of measure, 406
Fourier–Stieltjes coefficient, 349
Fourier–Stieltjes series, 349
frequently, 465
Fubini's Theorem, 15, 169, 170, 271, 273
Fubini's theorem on differentiation of series of monotone functions, 362
function, 556
functional, linear, 283
Fundamental Theorem of Algebra, 112, 572
Fundamental Theorem of Calculus, 35, 329, 363, 370

G_δ, 492
gamma function, 323, 353
gap, 484
geometric rectangle, 161, 266, 297
closed, 161
Gram determinant, 549
Gram matrix, 549
Gram–Schmidt orthogonalization process, 530, 549
graph, 291
greatest common divisor, 570

greatest lower bound, 003
Green's Theorem, 378

Hahn decomposition, 419
Hahn–Banach Theorem, 537
half space
  Poisson integral formula, 395
  Poisson kernel, 395
half-open interval, 3, 6, 451
Hardy–Littlewood maximal function, 330, 403, 438
Hardy–Littlewood Maximal Theorem, 328, 333, 430
harmonic function, 181, 334, 354, 392, 437
  unit disk, 227, 518
Hausdorff, 105, 447
Hausdorff–Young Theorem, 428
hedgehog space, 88
Heine–Borel Theorem, 108, 110
Helly's Selection Principle, 78
Helly–Bray Theorem, 405, 508
Herglotz's Theorem, 518
Hermitian inner product, 85, 89
Hermitian symmetric, 89
Hilbert cube, 88
Hilbert space, 523
  dimension, 534
Hilbert transform, 397, 433
  existence almost everywhere, 438
Hilbert–Schmidt norm, 138
Holder's inequality, 411
homeomorphism, 96, 443
homogeneous function, 180
homogeneous linear ordinary differential equation, 184, 201, 228

image, 556, 567
  direct, 559
  inverse, 559
imaginary part, 563
implicit differentiation, 153
Implicit Function Theorem, 155
indicator function, 171, 241
indicial equation, 222, 223
indiscrete space, 87
infimum, 3
infinite Taylor series, 44
infinitely differentiable, 46
infinity, 5, 6
  vanish at, 509
inhomogeneous linear ordinary differential equation, 201
initial condition, 186
inner measure, 255
inner product, 89, 523
inner-product space, 89
  pseudo, 90
integrable, 242, 276
  Lebesgue, 242
  Riemann, 27, 42, 162
  uniformly, 292
  vector-valued function, 274, 277
integral curve, 199
integral, Lebesgue, 242
integral, Riemann, 42
integration by parts, 36, 67, 344
interchange of limits, 13
interior, 92, 442
Interior Mapping Principle, 545
Intermediate Value Theorem, 12, 116
intersection, 555
interval, 3, 6
inverse function, 558
Inverse Function Theorem, 156, 562
inverse image, 559
irregular singular point, 227
isometric, 127
isometry, 127

Jacobian matrix, 140
Jessen–Marcinkiewicz–Zygmund, 438
Jordan and von Neumann Theorem, 526, 551
Jordan block, 212
Jordan decomposition, 418
Jordan form, 212
Jordan normal form, 213

kernel
  Dirichlet, 69
  Fejér, 71
  of linear function, 567

L^p completeness, 285, 414
L^p dual, 425
L^p norm, 410
L^p translation, 309, 417
Lagrange multipliers, 182, 411
Laplace equation, 181, 392

Laplacian, 181, 392
least upper bound, 2, 574
Lebesgue constant, 544
Lebesgue decomposition, 368, 423
Lebesgue integral, 242
Lebesgue measurable, 239, 262, 265, 305
Lebesgue measure, 118, 236, 298, 303
Lebesgue set, 371
Lebesgue's theorem on differentiation of monotone functions, 359
Legendre polynomial, 221
Legendre's equation, 220
Leibniz test, 19
length, 118
lexicographic ordering, 484
limit, 5, 463, 466
  Banach, 550
  of a sequence, 98
  point, 3, 93, 442
  uniform, 17
limits, interchange of, 13
Lindelöf space, 451
linear function, 566
linear functional, 283, 525
  norm, 424
  positive, 488
linear map, 566
linear operator, 283, 523
  bounded, 283, 523
  continuous, 283
linear ordinary differential equation, 184, 201
  constant coefficients, 208
  homogeneous, 228
linear transformation, 566
Lipschitz condition, 189
local base, 450
  countable, 450
localization of Fourier series, 71
localization of Lebesgue integral, 246
locally compact, 455
locally connected, 131, 483
locally finite open cover, 483
locally pathwise connected, 131, 483
logarithm, 49
long line, 484
lower bound, 3
  greatest, 3
lower Riemann integral, 27, 162
lower Riemann sum, 27, 162
lower semicontinuous, 481
lower-dimensional set, 325
Lusin's Theorem, 517

map, 556
  linear, 566
mapping, 556
Marcinkiewicz Interpolation Theorem, 430
matrix, 566
  elementary, 174
  Jacobian, 140
  Jordan form, 213
  Wronskian, 205
maximal element, 573
maximal function, 330
  Hardy–Littlewood, 330
Mean Value Theorem, 560
measurable, Borel, 239, 297
measurable function, 238
  Baire, 505
  Borel, 504
measurable, Lebesgue, 239, 262, 265, 305
measurable set, 238, 241, 257, 266
measurable vector-valued function, 275
measure, 235
  absolutely continuous, 420
  absolutely continuous Stieltjes, 369
  associated Stieltjes, 344
  Baire, 506
  Borel, 299, 316, 318, 490
  Cantor, 354, 519
  counting, 236
  delta, 342
  inner, 255
  Lebesgue, 236, 298, 303
  outer, 255
  product, 270
  purely irregular Borel, 519
  regular Borel, 299, 316, 318, 490
  regular Borel complex, 514
  regular Borel signed, 512
  signed, 418
  singular, 423
  singular Stieltjes, 368
  Stieltjes, 339
measure 0, 166, 265
measure space, 238
  completion, 262, 291
  finite, 238

σ-finite, 238
member, 555
mesh, 27, 161
metric, 83
discrete, 87
hedgehog, 88
product, 103
uniform, 87
metric space, 83
complete, 113
completion, 128, 133
connected, 115
metric subspace, 102, 104
metric topology, 442
metrizable, 476
Minkowski's inequality, 412
for integrals, 287, 415
monotone class, 268
Monotone Class Lemma, 268, 269
Monotone Convergence Theorem, 15, 250, 251
monotone increasing, 339
monotone sequence, 5
monotone set function, 494
multiplication formula, 375
multiplicative character, 407
multiplicity, 27
of a root, 572
multiplier, 384
multiplier operator, 384

negative, xxi
negative variation, 347
neighborhood, 442
of a point, 92
of a subset, 92
net, 465
universal, 468
Newton's method, 78
nonnegative set function, 234
norm, 90, 429, 521
convergence, 284
essential supremum, 280
Euclidean, 84, 85
Hilbert–Schmidt, 138
L^p, 410
of linear functional, 424
operator, 136, 283, 429, 523
supremum, 281
total-variation, 514
uniform, 281
weak-type, 430
normal, 105, 447
normed linear space, 281, 521
finite-dimensional, 521
pseudo, 280
weak topology, 444
weak-star topology on dual, 444
nowhere dense, 118
null space, 567
number
extended real, 6
real, 2

one-one, 558
one-point compactification, 456
onto, 558
open ball, 83
open cover, 106, 451
locally finite, 483
open interval, 3, 6
open mapping, 471
open neighborhood, 92, 442
open rectangle, 161
open set, 3, 83, 442
open subcover, 106, 451
open subspace, 102
open unit disk, 181
operator
bounded linear, 283
continuous linear, 283
linear, 283, 523
multiplier, 384
norm, 136, 283, 429, 523
self-adjoint, 536
sublinear, 429
order, 183, 188
order complete, 484
order topology, 484
ordered pair, 555
ordering
lexicographic, 484
partial, 573
simple, 573
total, 484, 573
ordinals, countable, 292
ordinary differential equation, 183
constant coefficients, 208
existence theorem, 189, 229

first-order linear, 228
homogeneous linear, 184, 201, 228
inhomogeneous linear, 201
linear, 184, 201
uniqueness theorem, 193
ordinary differential equations system, 188
with constant coefficients, 211
orthogonal, 89, 527
orthogonal complement, 528
orthogonal projection, 530
orthonormal, 529
orthonormal basis, 532
orthonormal set, maximal, 532
oscillation, 119, 167, 481
outer measure, 255

p-adic numbers, 133
pair
ordered, 555
unordered, 555
parallelogram law, 526
parameters, 198
Parseval's equality, 532
Parseval's Theorem, 74, 338
partial derivative, 139
partial ordering, 573
partition, 26, 161
partition of unity, 151, 483
path, 116, 199, 482
pathwise connected, 116, 482
locally, 131
periodic, 63
Picard iteration, 187
Picard–Lindelöf Existence Theorem, 189
Plancherel formula, 381
point, 555
point mass, 342
pointwise bounded, 22, 121, 478
Poisson integral, 355
conjugate, 398
formula for half space, 395
formula for unit disk, 354, 392, 437
Poisson kernel
conjugate, 398
for half space, 395
for unit disk, 392
Poisson Summation Formula, 389
polar coordinates, 173, 322
polarization, 527
polynomial
characteristic, 208, 214
in one indeterminate, 568
Legendre, 221
prime, 571
positive, xxi
positive definite function, 406
positive linear functional, 488
positive variation, 347
potential theory, 353
power series, 44, 46
prime polynomial, 571
primitive mapping, 175
probability, 171, 241, 339, 439, 486
product
Cartesian, 556, 557
dot, 84, 566
inner, 89, 523
measure, 270
metric, 103
topology, 443, 458
product of metric spaces, 103
product of sets, 556
product of σ-algebras, 266
projection, orthogonal, 530
Projection Theorem, 528
pseudo inner-product space, 90
pseudo normed linear space, 280
pseudometric, 83
pseudometric space, 83
pseudonorm, 90, 279
purely finitely additive, 438
purely irregular Borel measure, 519
Pythagorean Theorem, 526

quotient map, 445, 471
quotient space of a Banach space, 549
quotient topology, 443, 445, 471

radius of convergence, 45, 78
Radon–Nikodym Theorem, 421, 425, 439, 507, 515
range, 556
rank, 567
real number, 2
real part, 563
rearrangement, 12
rectangle, 161, 267, 266, 297
abstract, 266, 297

closed, 161
closed geometric, 161
geometric, 266, 297
open, 161
reduction of order, 228
refinement, 162
of partition, 28
reflexive, 564, 573
reflexive Banach space, 541
region under a graph, 291
regular, 105, 236, 447
regular Borel complex measure, 514
regular Borel measure, 299, 316, 318, 490
regular Borel signed measure, 512
regular singular point, 221
relation, 556
equivalence, 564
function, 556
set, 473
relative topology, 446
relatively dense, 131
restriction, 558
Riemann integrable, 27, 42, 162
Riemann integral, 27, 42, 162
Riemann sum, 27, 39, 42, 162, 163
Riemann zeta function, 390
Riemann–Lebesgue Lemma, 67, 335, 380
Riesz Convexity Theorem, 427
Riesz, Frigyes and Marcel, 427
Riesz Representation Theorem, 425, 490, 528
Riesz's Lemma, 358
Riesz–Fischer Theorem, 338, 383
ring of sets, 233
Rising Sun Lemma, 358
root, 568
Russell paradox, 553

saltus function, 366
Sard's Theorem, 324
sawtooth function, 65
scalar, 87, 279
scalar-valued nonnegative function, 274
Schroeder–Bernstein Theorem, 578
Schwartz function, 385
Schwartz space, 385
Schwarz inequality, 75, 84, 90, 526, 564
second axiom of countability, 450
second countable, 450
section, 267
self-adjoint, 536
semidefinite, 90
seminorm, 279
separable, 107, 446
separate points, 124, 462
separation of variables, 227
separation properties, 105
sequence, 4, 97
Cauchy, 9, 112
convergent, 5
doubly infinite, 5
monotone, 5
space, 87
series
binomial, 52, 56
Fourier (*see* Fourier series)
Fourier–Stieltjes, 349
power, 44, 46
Taylor, 46
trigonometric, 61
set, 554
set function, 234
additive, 234
bounded additive, 418
completely additive, 234
countably additive, 234
monotone, 494
nonnegative, 234
purely finitely additive, 438
σ-finite completely additive, 237
set theory, 553
Zermelo–Fraenkel, 554
side, 266
σ-algebra of sets, 233
σ-bounded set, 490
σ-compact, 457
σ-finite completely additive set function, 237
σ-finite measure space, 238
σ-ring of sets, 233
signed measure, 418
regular Borel, 512
Silverman–Toeplitz summability method, 80
simple function, 241
canonical expansion, 241
simple ordering, 573
sine, 47
singleton, 555
singular measure, 423
singular point

irregular, 227
regular, 221
singular Stieltjes measure, 368
smallest closed vector subspace, 281
smooth function, 142
smooth vector field, 199
solution, 183, 188
sphere, 126
area, 326
spherical coordinates, 325
standard basis, 566
standard Cantor set, 118
Stieltjes measure, 339
absolutely continuous, 369
associated, 344
singular, 368
Stone–Weierstrass Theorem, 124, 132, 480
strong differentiation, 431, 438
strong type, 429
subalgebra, 124
subcover, 451
subdivision point, 26
sublinear operator, 429
subnet, 468
subordinate to a cover, 151
subsequence, 5
subspace, 102, 104
closed, 102
metric, 102, 104
open, 102
smallest closed vector, 281
topological, 446
vector, 104
summability, 53
Abel, 54
Cesàro, 53
Silverman–Toeplitz, 80
support, 172, 300, 488
supremum, 2
essential, 280
norm, 281
symmetric, 89, 564
symmetric difference, 232
system of ordinary differential equations, 188
constant coefficients, 211
existence theorem, 189, 229
uniqueness theorem, 193

$\mathbf{T}_1$, 105, 447
Tauberian theorem, 55
Taylor series, 46
infinite, 44
Taylor's Theorem, 43, 146
tends, 5
Tietze Extension theorem, 483
topological space, 442
topological subspace, 446
topology, 442
discrete, 443
finest, 445
half-open interval, 451
metric, 442
order, 484
product, 443, 458
quotient, 443, 445, 471
relative, 446
upper, 481
weak, 443
total differential, 153
total ordering, 484, 573
total variation, 355, 514
total-variation norm, 514
totally bounded, 113
trace of matrix, 149
transfinite induction, 293
transitive, 564, 573
translation, 303
in L^p, 309, 417
triangle inequality, 41, 76, 83, 90, 279, 563
trigonometric series, 61
trivial ultrafilter, 294
Tychonoff Product Theorem, 460, 470
Tychonoff's Lemma, 452
type, 429
strong, 429
weak, 429

ultrafilter, 294
trivial, 294
ultrametric inequality, 133
unbounded interval, 3
Uniform Boundedness Theorem, 543
uniform Cauchy criterion, 18
uniform continuity, 111
uniform convergence, 17, 99
uniform limit, 17
uniform metric, 87
uniform norm, 281

uniformly bounded, 22, 99, 121, 478
uniformly Cauchy, 18
uniformly continuous function, 11
uniformly equicontinuous, 22, 121
uniformly integrable, 292
union, 555
unique factorization, 571
uniqueness theorem
  Fourier series, 73, 77, 338
  integral curves, 199
  system of ordinary differential equations, 193
unit disk, 126, 181
  harmonic function, 227, 518
  Poisson integral formula, 354, 392, 437
  Poisson kernel, 392
unit sphere, 126
universal net, 468
unordered pair, 555
upper bound, 2, 573
  least, 2
upper Riemann integral, 27, 162
upper Riemann sum, 27, 162
upper semicontinuous, 481
upper topology, 481
Urysohn Metrization Theorem, 476
Urysohn's Lemma, 474

vanish at infinity, 509
variation
  bounded, 346, 355
  negative, 347
  of parameters, 207
  positive, 347
  total, 355, 514
vector field, 199
vector subspace, 104
  smallest closed, 281
vector-valued function
  integrable, 274, 277
  measurable, 275

weak L^1, 330
weak topology, 443
  determined by functions, 443, 462
  on normed linear space, 444
weak type, 331, 429
weak-star convergence, 284, 355, 405, 437
weak-star topology of dual of normed linear space, 444
weak-type norm, 430
Weierstrass Approximation Theorem, 60
Weierstrass M test, 20
well ordered, 573, 576
Wiener's Covering Lemma, 331
Wronskian determinant, 205
Wronskian matrix, 205

Young's inequality, 428

Zermelo's well-ordering theorem, 576
Zermelo–Fraenkel set theory, 554
zeta function, 390
Zorn's lemma, 573