Graduate Texts in Mathematics 183

Springer
New York
Berlin
Heidelberg
Barcelona
Budapest
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Graduate Texts in Mathematics

(continued after index)

Robert E. Megginson

An Introduction to Banach Space Theory

Springer

Robert E. Megginson
Mathematics Department
3856 East Hall
University of Michigan
Ann Arbor, MI 48109-1109

Mathematics Subject Classification (1991): 46-01, 46Bxx, 47Axx

Library of Congress Cataloging-in-Publication Data
Megginson, Robert E.
An introduction to Banach space theory / Robert E. Megginson
p. cm. — (Graduate texts in mathematics ; 183)
Includes bibliographical references (p. –) and index.
ISBN 0-387-98431-3 (acid-free paper)
1. Banach spaces. I. Title. II. Series.
QA322.2.M44 1998
515'.732—dc21 97-52159

Printed on acid-free paper.

Reprinted in China by Beijing World Publishing Corporation, 2003

9 8 7 6 5 4 3 2 1

ISBN 0-387-98431-3 Springer-Verlag New York Berlin Heidelberg SPIN 10659704

To my mother and father
and, of course, to Kathy

Contents

*An asterisk preceding a section number indicates an optional section.

Preface

A *normed space* is a real or complex vector space X along with a *norm function* on the space; that is, a function $\|\cdot\|$ from X into the nonnegative reals such that if $x, y \in X$ and α is a scalar, then $\|\alpha x\| = |\alpha| \|x\|$, $\|x + y\| \leq \|x\| + \|y\|$, and $\|x\| = 0$ precisely when $x = 0$. A *Banach space* is then defined to be a normed space such that the metric given by the formula $d(x, y) = \|x - y\|$ is complete. In a sense, the study of Banach spaces is as old as the study of the properties of the absolute value function on the real numbers. However, the general theory of normed spaces and Banach spaces is a much more recent development. It was not until 1904 that Maurice René Frechet [80] suggested that it might be fruitful to extend the notion of limit from specific situations commonly studied in analysis to a more general setting. In his 1906 thesis [81], he developed the notion of a general metric space and immediately embarked on the study of real $C[a, b]$, the vector space of all real-valued continuous functions on a compact interval $[a, b]$ of the real line with the metric given by the formula $d(f, g) = \max\{\, |f(t) - g(t)| : t \in [a, b]\,\}$. In this seminal work on metric space theory, Fréchet was already emphasizing the important role played by the completeness of metrics such as that of $C[a, b]$. He was also doing a bit of Banach space theory since his metric for $C[a, b]$ is induced by a norm, as will be seen in Example 1.2.10 of this book. By 1908, Erhard Schmidt [211] was using the modern notation $\|x\|$ for the norm of an element x of ℓ_2, but the formulation of the general definition of a normed space had to wait a few more years.

Since the study of normed spaces for their own sake evolved rather than arose fully formed, there is some room to disagree about who founded the field. Albert Bennett came close to giving the definition of a normed space in a 1916 paper [23] on an extension of Newton's method for finding roots, and in 1918 Frédéric Riesz [195] based a generalization of the Fredholm theory of integral equations on the defining axioms of a complete normed space, though he did not use these axioms to study the general theory of such spaces. According to Jean Dieudonné [64], Riesz had at this time considered developing a general theory of complete normed spaces, but never published anything in this direction. In a paper that appeared in 1921, Eduard Helly [102] proved what is now called Helly's theorem for bounded linear functionals. Along the way, he developed some of the general theory of normed spaces, but only in the context of norms on subspaces of the vector space of all sequences of complex scalars.

The first undisputed efforts to develop the general theory of normed spaces appeared independently in a paper by Hans Hahn [98] and in Stefan Banach's thesis [10], both published in 1922. Both treatments considered only complete normed spaces. Though the growth of the general theory proceeded through the 1920s, the real impetus for the development of modern Banach space theory was the appearance in 1932 of Banach's book *Théorie des Opérations Linéaires* [13], which stood for years as the standard reference work in the field and is still profitable reading for the Banach space specialist today.

Many important reference works in the field have appeared since Banach's book, including, among others, those by Mahlon Day [56] and by Joram Lindenstrauss and Lior Tzafriri [156, 157]. While those works are classical starting points for the graduate student wishing to do research in Banach space theory, they can be formidable reading for the student who has just completed a course in measure theory, found the theory of L_p spaces fascinating, and would like to know more about Banach spaces in general.

The purpose of this book is to bridge that gap. Specifically, this book is for the student who has had enough analysis and measure theory to know the basic properties of the L_p spaces, and is designed to prepare such a student to read the type of work mentioned above as well as some of the current research in Banach space theory. In one sense, that makes this book a functional analysis text, and in fact many of the classical results of functional analysis are in here. However, those results will be applied almost exclusively to normed spaces in general and Banach spaces in particular, allowing a much more extensive development of that theory while placing correspondingly less emphasis on other topics that would appear in a traditional functional analysis text.

It should be made clear that this book is an introduction to the general theory of Banach spaces, not a detailed survey of the structure of the classical Banach spaces. Along the way, the reader will learn quite a bit about the classical Banach spaces from their extensive use in the theory, examples, and exercises. Those who find their appetite for those spaces whetted have an entire feast awaiting them in the volumes of Lindenstrauss and Tzafriri.

Prerequisites

Appendix A contains a detailed list of the prerequisites for reading this book. Actually, these prerequisites can be summarized very briefly: Anyone who has studied the first third of Walter Rudin's *Real and Complex Analysis* [202], which is to say the first six chapters of that book, will be able to read this book through, cover-to-cover, omitting nothing. Of course, this implies that the reader has had the basic grounding in undergraduate mathematics necessary to tackle Rudin's book, which should include a first course in linear algebra. Though some knowledge of elementary topology beyond the theory of metric spaces is assumed, the topology presented near the beginning of Rudin's book is enough. In short, all of this book is accessible to someone who has had a course in real and complex analysis that includes the duality between the Lebesgue spaces L_p and L_q when $1 < p < \infty$ and $p^{-1} + q^{-1} = 1$, as well as the Riesz representation theorem for bounded linear functionals on $C(K)$ where K is a compact Hausdorff space, and who has not slighted the usual prerequisites for such a course.

In fact, a large amount of this book is accessible at a much earlier stage in a student's mathematical career. The real reason for the measure theory prerequisite is to allow the reader to see applications of Banach space theory to the L_p spaces and spaces of measures, not because the measure theory is itself crucial to the development of the general Banach space theory in this book. It is quite possible to use this book as the basis for an undergraduate topics course in Banach space theory that concentrates on the metric theory of finite-dimensional Banach spaces and the spaces ℓ_p and c_0, for which the only prerequisites are a first course in linear algebra, a first course in real analysis without measure theory, and an introduction to metric spaces without the more general theory of topological spaces. Appendix A explains in detail how to do so. A list of the properties of metric spaces with which a student in such a course should be familiar can be found in Appendix B. Since the ℓ_p spaces are often treated in the main part of this book as special L_p spaces, Appendix C contains a development of the ℓ_p spaces from more basic principles for the reader not versed in the general theory of L_p spaces.

A Few Notes on the General Approach

The basic terminology and notation in this book are close to that used by Rudin in [200], with some extensions that follow the notation of Lindenstrauss and Tzafriri from [156] and [157].

Though most of the results in this text are divided in the usual way into lemmas, propositions, and theorems, a result that is really a theorem occasionally masquerades as an example. For instance, the theorems of Nikodým and Day that $(L_p[0,1])^* = \{0\}$ when $0 \le p < 1$ appear as Example 2.2.24, since such an L_p space is an example of a Hausdorff topological vector space whose dual does not separate the points of the space.

The theory in this book is developed for normed spaces over both the real and complex scalar fields. When a result holds for incomplete normed spaces as well as Banach spaces, the result is usually stated and proved in the more general form so that the reader will know where completeness is truly essential. However, results that can be extended from Banach spaces or arbitrary normed spaces to larger classes of topological vector spaces usually do not get the more general treatment unless the extension has a specific application to Banach space theory.

Any extensive treatment of the theory of normed spaces does require the study of two vector topologies that are not in general even metrizable, namely, the weak and weak* topologies. Much of the sequential topological intuition developed in the study of normed spaces can be extended to nonmetrizable vector topologies through the use of nets, so nets play a major role in many of the topological arguments given in this book. Since the reader might not be familiar with these objects, an extensive development of the theory of nets is given in the first section of Chapter 2.

This book is sprinkled liberally with examples, both to show the theory at work and to illustrate why certain hypotheses in theorems are necessary.

This book is also sprinkled liberally with historical notes and citations of original sources, with special attention given to mentioning dates within the body of the text so that the reader can get a feeling for the time frame within which the different parts of Banach space theory evolved. In ascribing credit for various results, I relied both on my own reading of the literature and on a number of other standard references to point me to the original sources. Among those standard references, I would particularly like to mention the excellent "Notes and Remarks" sections of Dunford and Schwartz's book *Linear Operators, Part I* [67], most of which are credited to Bob Bartle in the introduction to that work. Anyone interested in the rich history of this subject should read those sections in their entirety. I hasten to add that any error in attribution in the book you are holding is entirely mine.

In many cases, no citation of a source is given for a definition or result, especially when a result is very basic or a definition evolved in such a way that it is difficult to decide who should receive credit for it, as is the case for the definition of the norm function. It must be emphasized that *in no case does the lack of a citation imply any claim to priority on my part.*

The exercises, of which there are over 450, have several purposes. One obvious one is to provide the student with some practice in the use of the results developed in the text, and a few quite frankly have no reason for their existence beyond that. However, most do serve higher purposes. One is to extend the theory presented in the text. For example, Banach limits are defined and developed in Exercise 1.102. Another purpose of some of the exercises is to provide supplementary examples and counterexamples. Occasionally, an exercise presents an alternative development of a main result. For example, in Exercise 1.76 the reader is guided through Hahn's proof of the uniform boundedness principle, which is based on a gliding hump argument and does not use the Baire category theorem in any form. With the exception of a few extremely elementary facts presented in the first section of Chapter 1, none of the results stated and used in the body of the text have their proofs left as exercises. Very rarely, a portion of an example begun in the body of the text is finished in the exercises.

One final comment on the general approach involves the transliteration of Cyrillic names. I originally intended to use the modern scheme adopted by *Mathematical Reviews* in 1983. However, in the end I decided to write these names as the authors themselves did in papers published in Western languages, or as the names have commonly appeared in other sources. For example, the modern MR transliteration scheme would require that V. L. Šmulian's last name be written as Shmul'yan. However, Šmulian wrote many papers in Western languages, several of which are cited in this book, in which he gave his name the Czech diacritical transliteration that appears in this sentence. No doubt he was just following the custom of his time, but because of his own extensive use of the form Šmulian I have presented his name as he wrote it and would have recognized it.

Synopsis

Chapter 1 focuses on the metric theory of normed spaces. The first three sections present fundamental definitions and examples, as well as the most elementary properties of normed spaces such as the continuity of their vector space operations. The fourth section contains a short development of the most basic properties of bounded linear operators between normed spaces, including properties of normed space isomorphisms, which are then used to show that every finite-dimensional normed space is a Banach space.

The Baire category theorem for nonempty complete metric spaces is the subject of Section 1.5. This section is, in a sense, optional, since none of the results outside of optional sections of this book depend directly on it, though some such results do depend on a weak form of the Baire category theorem that will be mentioned in the next paragraph. However, this section has not been marked optional, since a student far enough along in his or her mathematical career to be reading this book should become familiar with Baire category. This section is placed just before the section on the open mapping theorem, closed graph theorem, and uniform boundedness principle for the benefit of the instructor wishing to substitute traditional Baire category proofs of those results for the ones given here.

Since a course in functional analysis is not a prerequisite for this book, the reader may not have seen the open mapping theorem, closed graph theorem, and uniform boundedness principle for Banach spaces. Section 1.6 is devoted to the development of those results. All three are based on a very specific and easily proved form of the Baire category theorem, presented in Section 1.3 as Theorem 1.3.14: Every closed, convex, absorbing subset of a Banach space includes a neighborhood of the origin.

In Section 1.7, the properties of quotient spaces formed from normed spaces are examined and the first isomorphism theorem for Banach spaces is proved: If T is a bounded linear operator from a Banach space X onto a Banach space Y, then Y and $X/\ker(T)$ are isomorphic as Banach spaces. Following a section devoted to direct sums of normed spaces, Section 1.9 presents the vector space and normed space versions of the Hahn-Banach extension theorem, along with their close relative, Helly's theorem for bounded linear functionals. The same section contains a development of Minkowski functionals and gives an example of how they are used to prove versions of the Hahn-Banach separation theorem. Section 1.10 introduces the dual space of a normed space, and has the characterizations up to isometric isomorphism of the duals of direct sums, quotient spaces, and subspaces of normed spaces. The next section discusses reflexivity and includes Pettis's theorem about the reflexivity of a closed subspace of a reflexive space and many of its consequences. Section 1.12, devoted to separability, includes the Banach-Mazur characterization of separable Banach spaces as isomorphs of quotient spaces of ℓ_1, and ends with the characterization of separable normed spaces as the normed spaces that are compactly generated so that the stage is set for the introduction of weakly compactly generated normed spaces in Section 2.8. This completes the basic material of Chapter 1.

The last section of Chapter 1, Section 1.13, is optional in the sense that none of the material in the rest of the book outside of other optional sections depends on it. This section contains a number of useful characterizations of reflexivity, including James's theorem. Some of the more basic of these are usually obtained as corollaries of the Eberlein-Šmulian theorem, but are included here since they can be proved fairly easily without it. The most important of these basic characterizations are repeated in Section 2.8 after the Eberlein-Šmulian theorem is proved, so this section can be skipped without fear of losing them. The heart of the section is a proof of the general case of James's theorem: A Banach space is reflexive if each bounded linear functional x^* on the space has the property that the supremum of $|x^*|$ on the closed unit ball of the space is attained somewhere on that ball. The proof given here is a detailed version of James's 1972 proof [117]. While the development leading up to the proof could be abbreviated slightly by delaying this section until the Eberlein-Šmulian theorem is available, there are two reasons for my not doing so. The first is that I wish to emphasize that the proof is really based only on the elementary metric theory of Banach spaces, not on arguments involving weak compactness, and the best way to do that is to give the proof before the weak topology has even been defined (though I do cheat a bit by defining weak sequential convergence without direct reference to the weak topology). The second is due to the reputation that James's theorem has acquired as being formidably deep. The proof is admittedly a bit intricate, but it is entirely elementary, not all that long, and contains some very nice ideas. By placing the proof as early as possible in this book, I hope to stress its elementary nature and dispel a bit of the notion that it is inaccessible.

Chapter 2 deals with the weak topology of a normed space and the weak* topology of its dual. The first section includes some topological preliminaries, but is devoted primarily to a fairly extensive development of the theory of nets, including characterizations of topological properties in terms of the accumulation and convergence of certain nets. Even a student with a solid first course in general topology may never have dealt with nets, so several examples are given to illustrate both their similarities to and differences from sequences. A motivation of the somewhat nonintuitive definition of a subnet is given, along with examples. The section includes a short discussion of topological groups, primarily to be able to obtain a characterization of relative compactness in topological groups in terms of the accumulation of nets that does not always hold in arbitrary topological spaces. Ultranets are not discussed in this section, since they are not really needed in the rest of this book, but a brief discussion of ultranets is given in Appendix D for use by the instructor who wishes to show how ultranets can be used to simplify certain compactness arguments.

Section 2.2 presents the basic properties of topological vector spaces and locally convex spaces needed for a study of the weak and weak* topologies. The section includes a brief introduction to the dual space of a topological vector space, and presents the versions of the Hahn-Banach separation theorem due to Mazur and Eidelheit as well as the consequences for locally convex spaces of Mazur's separation theorem that parallel the consequences for normed spaces of the normed space version of the Hahn-Banach extension theorem.

This is followed by a section on metrizable vector topologies. This section is marked optional since the topologies of main interest in this book are either induced by a norm or not compatible with any metric whatever. An F-space is defined in this section to be a topological vector space whose topology is compatible with a complete metric, without the requirement that the metric be invariant. Included is Victor Klee's result that every invariant metric inducing a topologically complete topology on a group is in fact a complete metric, which has the straightforward consequence that every F-space, as defined in this section, actually has its topology induced by a complete *invariant* metric, and thereby answers a question of Banach.

The versions of the open mapping theorem, closed graph theorem, and uniform boundedness principle valid for F-spaces are given in this section.

Section 2.4 develops the properties of topologies induced by families of functions, with special emphasis on the topology induced on a vector space X by a subspace of the vector space of all linear functionals on X.

The study of the weak topology of a normed space begins in earnest in Section 2.5. This section is devoted primarily to summarizing and extending the fundamental properties of this topology already developed in more general settings earlier in this chapter, and exploring the connections between the weak and norm topologies. Included is Mazur's theorem that the closure and weak closure of a convex subset of a normed space are the same. Weak sequential completeness, Schur's property, and the Radon-Riesz property are studied briefly.

Section 2.6 introduces the weak* topology of the dual space of a normed space. The main results of this section are the Banach-Alaoglu theorem and Goldstine's theorem. This is followed by a section on the bounded weak* topology of the dual space of a normed space, with the major result of this section being the Krein-Šmulian theorem on weakly* closed convex sets: A convex subset C of the dual space X^* of a Banach space X is weakly* closed if and only if the intersection of C with every positive scalar multiple of the closed unit ball of X^* is weakly* closed.

Weak compactness is studied in Section 2.8. It was necessary to delay this section until after Sections 2.6 and 2.7 so that several results about the weak* topology would be available. The Eberlein-Šmulian theorem is obtained in this section, as is the result due to Krein and Šmulian that the closed convex hull of a weakly compact subset of a Banach space is itself weakly compact. The corresponding theorem by Mazur on norm compactness is also obtained, since it is an easy consequence of the same lemma that contains the heart of the proof of the Krein-Šmulian result. A brief look is taken at weakly compactly generated normed spaces.

The goal of optional Section 2.9 is to obtain James's characterization of weakly compact subsets of a Banach space in terms of the behavior of bounded linear functionals. The section is relatively short since most of the work needed to obtain this result was done in the lemmas used to prove James's reflexivity theorem in Section 1.13.

The topic of Section 2.10 is extreme points of nonempty closed convex subsets of Hausdorff topological vector spaces. The Krein-Milman theorem is obtained, as is Milman's partial converse of that result.

Chapter 2 ends with an optional section on support points and subreflexivity. Included are the Bishop-Phelps theorems on the density of support points in the boundaries of closed convex subsets of Banach spaces and on the subreflexivity of every Banach space.

Chapter 3 contains a discussion of linear operators between normed spaces far more extensive than the brief introduction presented in Section 1.4. The first section of the chapter is devoted to adjoints of bounded linear operators between normed spaces. The second focuses on projections and complemented subspaces, and includes Whitley's short proof of Phillips's theorem that c_0 is not complemented in ℓ_∞.

Section 3.3 develops the elementary theory of Banach algebras and spectra, including the spectral radius formula, primarily to make this material available for the discussion of compact operators in the next section but also with an eye to the importance of this material in its own right.

Section 3.4 is about compact operators. Schauder's theorem relating the compactness of a bounded linear operator to that of its adjoint is presented, as is the characterization of operator compactness in terms of the bounded-weak*-to-norm continuity of the adjoint. Riesz's analysis of the spectrum of a compact operator is obtained, and the method used yields the result for real Banach spaces as well as complex ones. The Fredholm alternative is then obtained from this analysis. Much of the rest of the section is devoted to the approximation property, especially to Grothendieck's result that shows that the classical definition of the approximation property in terms of the approximability of compact operators by finite-rank operators is equivalent to the common modern definition in terms of the uniform approximability of the identity operator on compact sets by finite-rank operators. The section ends with a brief study of the relationship between Riesz's notion of operator compactness and Hilbert's property of complete continuity, and their equivalence for a linear operator whose domain is reflexive.

The final section of Chapter 3 is devoted to weakly compact operators. Gantmacher's theorem is obtained, as well as the equivalence of the weak compactness of a bounded linear operator to the weak*-to-weak continuity of its adjoint. The Dunford-Pettis property is examined briefly in this section.

The purpose of **Chapter 4** is to investigate Schauder bases for Banach spaces. The first section develops the elementary properties of Schauder bases and presents several classical examples, including Schauder's basis for $C[0,1]$ and the Haar basis for $L_p[0,1]$ when $1 \leq p < \infty$. Monotone bases and the existence of basic sequences are covered, and the relationship between Schauder bases and the approximation property is discussed.

Unconditional bases are investigated in Section 4.2. Results are presented about equivalently renorming Banach spaces with unconditional bases to be Banach algebras and Banach lattices. It is shown that neither the classical Schauder basis for $C[0,1]$ nor the Haar basis for $L_1[0,1]$ is unconditional.

Section 4.3 is devoted to the notion of equivalent bases and applications to finding isomorphic copies of Banach spaces inside other Banach spaces. Characterizations of the standard unit vector bases for c_0 and ℓ_1 are given. Weakly unconditionally Cauchy series are examined, and the Orlicz-Pettis theorem and Bessaga-Pełczyński selection principle are obtained.

The properties of the sequence of coordinate functionals for a Schauder basis are taken up in Section 4.4, and shrinking and boundedly complete bases are studied. The final section of Chapter 4 is optional and is devoted to an investigation of James's space J, which was the first example of a nonreflexive Banach space isometrically isomorphic to its second dual.

Chapter 5 focuses on various forms of rotundity, also called strict convexity, and smoothness. The first section of the chapter is devoted to characterizations of rotundity, its fundamental properties, and examples, including one due to Klee that shows that rotundity is not always inherited by quotient spaces. The next section treats uniform rotundity, and includes the Milman-Pettis theorem as well as Clarkson's theorem that the L_p spaces such that $1 < p < \infty$ are uniformly rotund. Section 5.3 is devoted to generalizations of uniform rotundity, and discusses local uniform rotundity, weak uniform rotundity, weak* uniform rotundity, weak local uniform rotundity, strong rotundity, and midpoint local uniform rotundity, as well as the relationships between these properties.

The second half of Chapter 5 deals with smoothness. Simple smoothness is taken up in Section 5.4, in which the property is defined in terms of the uniqueness of support hyperplanes for the closed unit ball at points of the unit sphere and then characterized by the Gateaux differentiability of the norm and in several other ways. The partial duality between rotundity and smoothness is examined, and other important properties of smoothness are developed. Uniform smoothness is the subject of the next section, in which the property is defined using the modulus of smoothness and characterized in terms of the uniform Fréchet differentiability of the norm. The complete duality between uniform smoothness and uniform rotundity is proved. Fréchet smoothness and uniform Gateaux smoothness are examined in the final section of the chapter, and Šmulian's results on the duality between these properties and various generalizations of uniform rotundity are obtained.

Appendix A includes an extended description of the prerequisites for reading this book, along with a very detailed list of the changes that must be made to the presentation in Chapter 1 if this book is to be used for an undergraduate topics course in Banach space theory. **Appendices B** and **C** are included to support such a topics course. They are, respectively, a list of the properties of metric spaces that should be familiar to a student in such a course and a development of ℓ_p spaces from basic principles of analysis that does not depend on the theory of L_p spaces. **Appendix D** is a discussion of ultranets that supplements the material on nets in Section 2.1.

Dependences

No material in any nonoptional section of this book depends on material in any optional section, with the exception of a few exercises in which the dependence is clearly indicated. Where an optional section depends on other optional sections, that dependence is stated clearly at the beginning of the section.

The material in the nonoptional sections of Chapters 1 through 3 is meant to be taken up in the order presented, and each such section should be considered to depend on every other nonoptional section that precedes it. One important exception is that, as has already been mentioned, Section 1.5 can be omitted, since its results are used only in optional Section 2.3. However, the reader unfamiliar with the Baire category theorem will not want to skip this material.

All of the nonoptional sections of Chapters 1 and 2 should be covered before taking up Chapters 4 and 5. Chapter 4 also depends on the first two sections of Chapter 3. If the small amount of material in Chapter 4 on the approximation property is not to be skipped, then the development of that property in Section 3.4 must also be covered.

Some results about adjoint operators from Section 3.1 are used in Example 5.4.13. Except for this, Chapter 5 does not depend on the material in Chapters 3 and 4.

Appendix A does not depend on any other part of this book, except where it refers to changes that must be made to the presentation in Chapter 1 for an undergraduate topics course. Appendices B and C do not use material from any other portion of this book. Appendix D depends on Section 2.1 but on no other part of the book.

Acknowledgments

There are many who contributed valuable suggestions and various forms of aid and encouragement to this project. Among those, I would particularly like to mention Sheldon Axler, Mahlon Day, Alphonso DiPietro, Fred Gehring, John LeDuc, Don Lewis, Tenney Peck, M. S. Ramanujan, Ira Rosenholtz, Mark Smith, B. A. Taylor, and Jerry Uhl. I would also like to express my appreciation for the support and editorial assistance I received from Springer-Verlag, particularly from Ina Lindemann and Steve Pisano. This book would hardly have been possible without the day-to-day support of my wife Kathy, who did a wonderful job of insulating me from the outside world during its writing and who did not hesitate to pass stern judgment on some of my more convoluted prose. Finally, my special thanks go to my friend and mentor Horacio Porta, who read an early version of the manuscript and made some extremely valuable suggestions that substantially improved the content and format of this book.

Ann Arbor
September, 1998

ROBERT E. MEGGINSON

1

Basic Concepts

This chapter contains the basic definitions and initial results needed for a study of Banach spaces. In particular, the material presented in the first twelve sections of this chapter, with the exception of that of Section 1.5, is used extensively throughout the rest of this book. Section 1.13, though containing material that is very important in modern Banach space theory, is optional in the sense that the few results and exercises in the rest of the book that depend on this material are clearly marked as such.

1.1 Preliminaries

Here are some of the definitions, conventions, and notation used throughout this book. Whenever a definition contains two or more different names for the same object, the first is the one usually used here. The alternative names are included because they are sometimes encountered in other sources.

The set of positive integers is denoted by $\mathbb{N}$. The fields of real and complex numbers are denoted by $\mathbb{R}$ and $\mathbb{C}$ respectively. The symbol $\mathbb{F}$ denotes a field that can be either $\mathbb{R}$ or $\mathbb{C}$. The elements of $\mathbb{F}$ are called *scalars.*

In a topological space, the *closure* of a set A, denoted by $\overline{A}$, is the smallest closed set that includes A, that is, the intersection of all closed sets that include A. The *interior* of A, denoted by A°, is the largest open subset of A, that is, the union of all open subsets of A.

A *vector space* or *linear space* over $\mathbb{F}$ is a set X of objects called *vectors* along with an operation $+$ from $X \times X$ into X called *addition of vectors* and an operation $\cdot$ from $\mathbb{F} \times X$ into X called *multiplication of vectors by scalars* satisfying these conditions:

(1) addition is commutative and associative;

(2) there is a *zero vector* 0 in X, sometimes called the *origin* of X, such that $x + 0 = x$ for each vector x;

(3) for each vector x there is a vector $-x$ such that $x + (-x) = 0$;

(4) for all scalars α and β and all vectors x and y, $\alpha \cdot (x+y) = \alpha \cdot x + \alpha \cdot y$, $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$, and $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$;

(5) for each vector x, $1 \cdot x = x$.

The *difference* $x - y$ of vectors x and y is the vector $x + (-y)$. Scalars are usually represented in this book by lowercase letters near the beginning of the Greek alphabet and vectors by lowercase letters near the end of the Roman alphabet, with one major exception being that nonnegative real numbers are often denoted by the letters r, s, and t. A product $\alpha \cdot x$ of a scalar and a vector is usually abbreviated to αx. Though the symbol 0 is used for the zeros of both $\mathbb{F}$ and X, the context should always make it clear which is intended. In this book, a *subspace* of a vector space X always means a *vector* subspace, that is, a subset of X that is itself a vector space under the same operations. It is presumed that the reader is familiar with linear independence, bases, and other elementary vector space concepts.

It is worth emphasizing that *in this book the term "vector space" always means a vector space over* $\mathbb{R}$ *or* $\mathbb{C}$. The terms *real vector space* and *complex vector space* are used when it is necessary to be specific about the scalar field. Another important convention is that *except where stated otherwise, all vector spaces discussed within the same context are assumed to be over the same field* $\mathbb{F}$, *but* $\mathbb{F}$ *may be either* $\mathbb{R}$ *or* $\mathbb{C}$. For example, suppose that a theorem begins with the sentence "Let X be a normed space and Y a Banach space" and that no field is mentioned anywhere in the theorem. Normed spaces and Banach spaces, defined in the next section, are vector spaces with some additional structure. Thus, it is implied that X and Y are either both real or both complex.

Notational difficulties can arise when a subset of a vector space is linearly independent, but a particular description of the set in an indexed form such as $\{ x_\alpha : \alpha \in I \}$ allows x_{α_1} and x_{α_2} to be equal when $\alpha_1 \neq \alpha_2$. In every instance in this book that such notation could cause a problem, the set turns out to be countable,[1] and the difficulty is avoided by writing the set as a finite list or a sequence. A finite list $x_1, \ldots, x_{n_0}$ or a sequence (x_n) of elements in a vector space is said to be linearly independent if the corresponding set $\{x_1, \ldots, x_{n_0}\}$ or $\{ x_n : n \in \mathbb{N} \}$ is linearly independent and $x_j \neq x_k$ when $j \neq k$.

---

[1] In this book, finite sets as well as countably infinite ones are said to be countable.

Notice that the sequence in the preceding paragraph was assumed to be indexed by $\mathbb{N}$. *Except where stated otherwise, sequences are indexed by the positive integers in their natural order; in particular, the indexing starts with* 1 *rather than* 0. This will be important when quantities related to the terms of a sequence (x_n) are stated as functions of the index value n.

Suppose that X is a vector space, that x is an element of X, that A and B are subsets of X, and that α is a scalar. Then

$$\begin{aligned} x + A &= \{\, x + y : y \in A \,\}, \\ x - A &= \{\, x - y : y \in A \,\}, \\ A + B &= \{\, y + z : y \in A,\ z \in B \,\}, \\ A - B &= \{\, y - z : y \in A,\ z \in B \,\}, \\ \alpha A &= \{\, \alpha y : y \in A \,\}, \text{ and} \\ -A &= \{\, -y : y \in A \,\}. \end{aligned}$$

The set $x+A$ is called the *translate* of A by x. Notice that $A-B$ represents the *algebraic* difference of the sets A and B. The *set-theoretic* difference $\{\, x : x \in A,\ x \notin B \,\}$ is denoted by $A \setminus B$.

Suppose that A is a subset of a vector space X. Then A is *convex* if $ty + (1-t)z \in A$ whenever $y, z \in A$ and $0 < t < 1$. If $\alpha A \subseteq A$ whenever $|\alpha| \leq 1$, then A is *balanced*. The set A is *absorbing* if, for each x in X, there is a positive number s_x such that $x \in tA$ whenever $t > s_x$. The following properties of these special types of sets are not difficult to prove; see Exercises 1.1 and 1.3.

(1) Absorbing sets always contain 0. So do *nonempty* balanced sets.

(2) If A is a balanced set, then $\alpha A = A$ whenever $|\alpha| = 1$, which in particular implies that $-A = A$.

(3) Arbitrary unions and intersections of balanced sets are balanced.

(4) Arbitrary intersections of convex sets are convex.

(5) Translates and scalar multiples of convex sets are convex.

(6) The set A is convex if and only if $sA+tA = (s+t)A$ whenever $s, t > 0$.

The *convex hull* or *convex span* of A, denoted by $\operatorname{co}(A)$, is the smallest convex set that includes A, that is, the intersection of all convex sets that include A. It is not difficult to show that $\operatorname{co}(A)$ is the collection of all *convex combinations* of elements of A, that is, all sums of the form $\sum_{j=1}^{n} t_j x_j$ such that $n \in \mathbb{N}$, $x_1, \ldots, x_n \in A$, $t_1, \ldots, t_n \geq 0$, and $\sum_{j=1}^{n} t_j = 1$. See Exercise 1.4. If X has a topology, then the *closed convex hull* or *closed convex span* of A, denoted by $\overline{\operatorname{co}}(A)$, is the smallest closed convex set that includes A, that is, the intersection of all closed convex sets that include A. The *linear hull* or *linear span* of A, denoted by $\langle A \rangle$, is the smallest subspace of X that includes A, that is, the intersection of all subspaces of X that include A. Notice that $\langle \varnothing \rangle = \{0\}$. If A is nonempty, then it is not difficult to show that $\langle A \rangle$ is the collection of all *linear combinations* of elements of A, that is, all sums of the form $\sum_{j=1}^{n} \alpha_j x_j$ such that $n \in \mathbb{N}$, $x_1, \ldots, x_n \in A$, and $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. See Exercise 1.4. If X has a topology, then the *closed linear hull* or *closed linear span* of A, denoted by $[A]$, is the smallest closed subspace of X that includes A, that is, the intersection of all closed subspaces of X that include A. See Exercise 1.12 for an important observation about $\overline{\text{co}}(A)$ and $[A]$.

Two conventions are important for the interpretation of statements about arbitrary unions and intersections, such as statements (3) and (4) in the preceding paragraph. The first is that *the union of an empty family of sets is the empty set,* while the second is that *the intersection of an empty family of sets is the universal set from which subsets are being taken.* Both "conventions" actually follow directly from rigorous applications of the definitions of arbitrary unions and intersections of sets; see, for example, [65]. However, some authors do prefer to leave intersections of empty families undefined because of the confusion that occurs when it is not clear what set is the universal set. The context will almost always prevent that problem from arising here; for example, in the preceding paragraph it is clear that the universal set for the intersections in statements (3) and (4) is the vector space X. Where confusion might otherwise occur, it will be made clear what set is considered universal. In any case, every statement in this book about a union or intersection of an arbitrary family of subsets of some universal set is intended to apply also to the empty family, which should be kept in mind when interpreting such a statement.

Let X and Y be vector spaces. A *linear operator* or *linear function* or *linear transformation* from X into Y is a function $T: X \to Y$ such that the following two conditions are satisfied whenever $x, x_1, x_2 \in X$ and $\alpha \in \mathbb{F}$:

(1) $T(x_1 + x_2) = T(x_1) + T(x_2)$;

(2) $T(\alpha x) = \alpha T(x)$.

If $\mathbb{F}$ is viewed as a one-dimensional vector space, then a linear operator from X into $\mathbb{F}$ is called a *linear functional* or *linear form* on X. For linear mappings only, the notation Tx is often used as an abbreviation for $T(x)$. The *kernel* or *null space* of a linear operator T, denoted by $\ker(T)$, is the subspace $T^{-1}(\{0\})$ of X. That is, $\ker(T) = \{\, x : x \in X,\, Tx = 0 \,\}$. Notice that a linear operator T is one-to-one if and only if $\ker(T) = \{0\}$. The *rank* of a linear operator is the dimension of its range. Thus, a *finite-rank* linear operator is a linear operator with a finite-dimensional range. If S is a linear operator from X into Y and T is a linear operator from Y into a vector space Z, then the *product* or *composite*[2] TS of S and T is the

[2]Of course, the composite $g \circ f$ is defined for any functions f and g, linear or not, such that the range of f lies in the domain of g, but in the nonlinear case it is customary to insert the $\circ$ symbol and not to use the term *product*, especially in situations in which composites could be confused with pointwise products of scalar-valued functions.

linear operator from X into Z formed by letting $TS(x) = T(S(x))$ for each x in X. Linear operators preserve some of the special properties of sets mentioned above. The following are easy to show; see Exercise 1.11.

(1) Let A be a subset of the vector space X and let T be a linear operator with domain X. If A is convex, or balanced, or a subspace, then $T(A)$ has the same property.

(2) Let B be a subset of the vector space Y and let T be a linear operator with range in Y. If B is convex, or balanced, or absorbing, or a subspace, then $T^{-1}(B)$ has the same property.

It is presumed that the reader is familiar with standard facts about linear operators, including the fact that the collection of all linear operators from X into Y is a vector space with the vector space operations given by the usual addition of functions and multiplication of functions by scalars. This vector space of linear operators is denoted by $L(X,Y)$. The vector space $L(X,\mathbb{F})$ of all linear functionals on X is denoted by $X^{\#}$.

A *metric space* is a set M along with a *metric* or *distance function* $d\colon M \times M \to \mathbb{R}$ such that the following three conditions are satisfied by all x, y, and z in M:

(1) $d(x,y) \geq 0$, and $d(x,y) = 0$ if and only if $x = y$;

(2) $d(x,y) = d(y,x)$;

(3) $d(x,z) \leq d(x,y) + d(y,z)$ (the triangle inequality).

If $x \in M$ and $r > 0$, then the *closed ball* in M with center x and radius r is $\{\, y : y \in M,\ d(x,y) \leq r \,\}$ and is denoted by $B(x,r)$. The corresponding *open ball* is $\{\, y : y \in M,\ d(x,y) < r \,\}$. Appendix B has a list of some elementary metric space properties with which the reader should be familiar.

Suppose that M is a metric space with metric d and that A is a nonempty subset of M. If B is another nonempty subset of M, then the *distance $d(A,B)$ from A to B* is $\inf\{\, d(x,y) : x \in A,\ y \in B \,\}$. For each x in M, the *distance $d(x,A)$ from x to A* is $\inf\{\, d(x,y) : y \in A \,\}$. It is not hard to show that $|d(x,A) - d(y,A)| \leq d(x,y)$ for all x and y in M, from which it follows immediately that the function $x \mapsto d(x,A)$ is continuous on M. See Exercise 1.13.

A *preorder* on a set A is a binary relation $\preceq$ on A satisfying the following two conditions:

(1) $a \preceq a$ for each a in A;

(2) if $a \preceq b$ and $b \preceq c$, then $a \preceq c$.

A set with a preorder is called a *preordered* set. Suppose that A is a preordered set. A *chain* in A is a subset C of A such that for all a and b in C, either $a \preceq b$ or $b \preceq a$. An *upper bound* for a subset B of A is an element u of A such that $b \preceq u$ for each b in B. A *maximal element* in A is an element m of A such that whenever $m' \in A$ and $m \preceq m'$, then $m' \preceq m$.

The axiom of choice is going to be used frequently and fearlessly, often in the following form.

Zorn's Lemma. (M. Zorn, 1935 [249]). *A preordered set in which each chain has an upper bound contains at least one maximal element.*

A proof that Zorn's lemma is equivalent as an axiom of set theory to the axiom of choice can be found in [65]. The hard part of that proof, involving a "tower" argument to show that the axiom of choice implies Zorn's lemma, can be replaced by a more recent and greatly simplified argument due to Jonathan Lewin [152].

Zorn's lemma is sometimes stated only for *partially ordered* sets, that is, preordered sets also having this property:

(3) if $a \preceq b$ and $b \preceq a$, then $a = b$.

As axioms of set theory, these two forms of Zorn's lemma are equivalent; see Exercise 1.16.

The following is a typical application of Zorn's lemma.

1.1.1 Theorem. *If X is a vector space and S is a linearly independent subset of X, then X has a basis that includes S. In particular, every vector space has a basis.*

PROOF. Let the collection of all linearly independent subsets of X that include S be denoted by $\mathfrak{A}$.[3] Preorder the members of $\mathfrak{A}$ by inclusion; that is, define a preorder $\preceq$ on $\mathfrak{A}$ by declaring that $A_1 \preceq A_2$ when $A_1 \subseteq A_2$. If $\mathfrak{C}$ is a chain in $\mathfrak{A}$, then $S \cup \{\bigcup\{C : C \in \mathfrak{C}\}\}$ is an upper bound for $\mathfrak{C}$ in $\mathfrak{A}$. By Zorn's lemma, the collection $\mathfrak{A}$ must have a maximal element M. If M were not a basis for X, then there would be an x in $X \setminus \langle M \rangle$, and $M \cup \{x\}$ would be a member of $\mathfrak{A}$ with the property that $M \preceq M \cup \{x\}$ but $M \cup \{x\} \not\preceq M$, a contradiction to the maximality of M.

The assertion that each vector space has a basis now follows, since each vector space must have a basis that includes the linearly independent set $\emptyset$. ■

Exercises

Unless stated otherwise, in each of these exercises X is a vector space, the elements x and y are vectors in X, the sets A, B, and C are subsets of X, and α and β are scalars.

---

[3] The symbol $\mathfrak{A}$ is the Fraktur letter A. Uppercase Fraktur letters $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \ldots$ will often be used to denote certain types of sets, particularly those whose members are sets or functions. See the List of Symbols for the list of Fraktur letters corresponding to the uppercase Roman letters.

1.1 (a) Show that if A is an absorbing set or a nonempty balanced set, then $0 \in A$.

(b) Show that if A is balanced, then $\alpha A = A$ whenever $|\alpha| = 1$.

(c) Suppose that $\mathfrak{B}$ is a collection of balanced subsets of X. Show that $\bigcup\{S : S \in \mathfrak{B}\}$ and $\bigcap\{S : S \in \mathfrak{B}\}$ are both balanced.

(d) Suppose that $\mathfrak{C}$ is a collection of convex subsets of X. Show that $\bigcap\{S : S \in \mathfrak{C}\}$ is convex.

(e) Show that if A is convex, then $x + A$ and αA are convex.

1.2 (a) Show that the "addition" and "multiplication by scalars" defined for sets obey the commutative and associative laws for vector spaces. That is, show that $A+B = B+A$, that $A+(B+C) = (A+B)+C$, and that $\alpha(\beta A) = (\alpha\beta)A$. Show also that $(x + A) + (y + B) = (x + y) + (A + B)$.

(b) Show that $\alpha(A + B) = \alpha A + \alpha B$.

(c) Show that $(\alpha + \beta)A \subseteq \alpha A + \beta A$ but that equality need not hold.

1.3 (a) Prove that A is convex if and only if $sA + tA = (s + t)A$ for all positive s and t. (Consider the special case in which $s + t = 1$.)

(b) Use (a) and Exercise 1.2 to prove that if A and B are convex, then so is $A + B$.

1.4 Show that the convex hull of A is the collection of all convex combinations of elements of A. Show that if A is nonempty, then the linear hull of A is the collection of all linear combinations of elements of A.

1.5 Prove that if A and B are balanced, then so is $A + B$.

1.6 Suppose that A is balanced. Prove that A is absorbing if and only if the following holds: For each x in X there is a positive number t_x such that $x \in t_x A$.

1.7 Identify all of the balanced subsets of $\mathbb{C}$. Do the same for $\mathbb{R}^2$.

1.8 The *balanced hull* $\mathrm{bal}(A)$ of A is the smallest balanced subset of X that includes A. There is a simple expression for $\mathrm{bal}(A)$ as the union of a certain collection of sets. Find it.

1.9 Prove that each convex absorbing subset of $\mathbb{C}$ includes a neighborhood of 0.

1.10 Is the conclusion of the preceding exercise true for nonconvex absorbing subsets of $\mathbb{C}$?

1.11 Let X and Y be vector spaces. Suppose that $A \subseteq X$, that $B \subseteq Y$, and that $T\colon X \to Y$ is a linear operator.

(a) Prove that if B is convex, or balanced, or absorbing, or a subspace, then $T^{-1}(B)$ has that same property.

(b) Prove that if A is convex, or balanced, or a subspace, then $T(A)$ has that same property. Prove that if T maps X onto Y and A is absorbing, then $T(A)$ is absorbing.

1.12 Let X be the vector space $\mathbb{R}^2$ with the topology whose only member besides the entire space and the empty set is $\{(\alpha,\beta) : \beta > \alpha^2\}$. Find a subspace A of X such that $\overline{A}$ is not convex (and therefore is not a subspace of X), so that neither of the equations $[A] = \overline{\langle A \rangle}$ and $\overline{\text{co}}(A) = \overline{\text{co}(A)}$ holds. (It will be shown in Sections 1.3 and 2.2 that both equations do hold for every subset A of a topologized vector space under certain restrictions on the topology, in particular when it comes from a norm.)

1.13 Let M be a metric space with metric d. Suppose that A is a nonempty subset of M and that x and y are elements of M. Show that $d(x,A) \leq d(x,y) + d(y,A)$. Conclude that $|d(x,A) - d(y,A)| \leq d(x,y)$.

1.14 Let $\mathfrak{C}$ be the collection of nonempty closed bounded subsets of a metric space M. The *Hausdorff distance* $d_H(A,B)$ between two members A and B of $\mathfrak{C}$ is defined as follows. Let $\rho(A,B) = \sup\{d(x,B) : x \in A\}$, and then let $d_H(A,B) = \max\{\rho(A,B), \rho(B,A)\}$. Prove that d_H is a metric on $\mathfrak{C}$.

1.15 Prove that if S is a subset of a vector space X, then S has a subset B such that B is a basis for $\langle S \rangle$. Use this to give another proof that every vector space has a basis.

1.16 (a) Suppose that P is a preordered set. Define a binary relation $\sim$ on P by declaring that $x \sim y$ when $x \preceq y$ and $y \preceq x$. Prove that $\sim$ is an equivalence relation. Now let $P/\!\sim$ be the collection of equivalence classes determined by $\sim$. For each x in P, let $[x]$ represent the equivalence class containing x. Declare that $[x] \preceq [y]$ when $x \preceq y$. Show that this unambiguously defines a relation on $P/\!\sim$. Show that this relation partially orders $P/\!\sim$.

(b) Use (a) to derive the form of Zorn's lemma for preordered sets from the form for partially ordered sets. Conclude that the two forms of Zorn's lemma are equivalent as axioms of set theory.

1.17 Prove that the following principle is equivalent to Zorn's lemma as an axiom of set theory.

Hausdorff's Maximal Principle. (F. Hausdorff, 1914 [101, p. 140]). *Every preordered set includes a chain that is not a proper subset of another chain.*

1.2 Norms

Suppose that X is a vector space. For each ordered pair (x_1, x_2) of elements of X, define $(x_1, x_2]$ to be $\{y : y = (1-t)x_1 + tx_2,\ 0 < t \leq 1\}$. In particular, this implies that $(x,x] = \{x\}$ for each x in X. It is not difficult to check that if $(x_1,x_2] = (y_1,y_2]$, then $x_1 = y_1$ and $x_2 = y_2$; see Exercise 1.18. Define an equivalence relation on the collection of all of these "half-open line segments" in X by declaring that $(x_1,x_2]$ is equivalent to $(y_1,y_2]$ whenever there is a z in X such that $y_1 = x_1 + z$ and $y_2 = x_2 + z$. That is, two of these segments are equivalent when one is a translate of the other. Let the *arrow* $\overrightarrow{x_1, x_2}$ *with head* x_2 *and tail* x_1 be the equivalence class containing $(x_1, x_2]$. It is easy to see that the collection X_a of all such arrows is a vector space over $\mathbb{F}$ when given these operations:

$$\overrightarrow{x_1, x_2} + \overrightarrow{y_1, y_2} = \overrightarrow{x_1 + y_1, x_2 + y_2},$$
$$\alpha \cdot \overrightarrow{x_1, x_2} = \overrightarrow{\alpha x_1, \alpha x_2}.$$

It is also easy to see that the map $x \mapsto \overrightarrow{0, x}$ is a vector space isomorphism from X onto X_a. Thus, the vectors of X can be regarded either as its elements or as directed line segments in X that remain fundamentally unchanged when translated. In fact, the familiar "arrow vectors" of calculus are obtained by treating the vectors of $\mathbb{R}^2$ or $\mathbb{R}^3$ the second way.

The fundamental metric notion for vectors can be either distance or length, depending on the way vectors are treated. If they are thought of as points, then the *distance* between two vectors is a reasonable notion, while the *length* of a vector would seem to be a concept best left undefined. On the other hand, the length of a vector is an entirely natural concept if vectors are considered to be arrows. The distance between two vectors would then seem to be a more nebulous notion, though it might be reasonable to obtain it, as is done with the arrow vectors of calculus, by joining the tails of the two vectors and then measuring the length of one of the two "difference vectors" between their heads.

The concept of a *norm* comes from thinking of vectors as arrows. A norm on a vector space is a function that assigns to each vector a length. There are some obvious properties that such a function should be required to have. A nonzero vector should have positive length; the additive inverse of a vector should have the same length as the original vector; half a vector should have half the length of the full vector; a vector forming one side of a triangle should not be longer than the sums of the lengths of the vectors forming the other two sides; and so forth. These requirements are embodied in the following definition.

1.2.1 Definition. Let X be a vector space. A *norm* on X is a real-valued function $\|\cdot\|$ on X such that the following conditions are satisfied by all members x and y of X and each scalar α:

(1) $\|x\| \geq 0$, and $\|x\| = 0$ if and only if $x = 0$;

(2) $\|\alpha x\| = |\alpha|\,\|x\|$;

(3) $\|x + y\| \leq \|x\| + \|y\|$ (the triangle inequality).

The ordered pair $(X, \|\cdot\|)$ is called a *normed space* or *normed vector space* or *normed linear space*.

The term *norm* is commonly used for both the function of the preceding definition and its values, so that $\|x\|$ is read as "the norm of x." Another common practice is to reserve the rigorous ordered-pair notation of the definition for situations in which several different norms have been defined on the same vector space, making it necessary to be able to distinguish between the resulting normed spaces. When there is no danger of confusion, it is customary to use the same symbol, such as X, to denote the normed space, the vector space underlying the normed space, and the set underlying the vector space. Corresponding shortcuts are often used in other language referring to normed spaces, for example by saying that some object is an element of a normed space rather than an element of the set underlying the vector space underlying the normed space. Notice that this is a straightforward extension of the common practice of having the same symbol represent a vector space and the set underlying it, rather than using a symbol such as X to denote the underlying set and then referring to the vector space as an ordered triple $(X, +, \cdot)$ formed from the set and the two vector space operations.

The distance between two vectors x and y in a normed space X can be defined just as it is for the arrow vectors of calculus; that is, the distance between x and y is the length of the difference vector $x - y$. It is easy to check that this does in fact give a metric on X.

1.2.2 Definition. Let X be a normed space. The *metric induced by the norm of* X is the metric d on X defined by the formula $d(x, y) = \|x - y\|$. The *norm topology of* X is the topology obtained from this metric.

As will be seen in the next chapter, there are other natural topologies that normed spaces possess besides their norm topologies. *Henceforth, whenever reference is made to some topological property such as compactness or convergence in a normed space without specifying the topology, the norm topology is implied.*

Suppose that X is a normed space and that $B(x, r)$ is the closed ball centered at a point x of X and having radius r. It follows easily from the definition of the metric of X that $B(x, r) = x + rB(0, 1)$. A corresponding relationship exists between open balls in X and the open ball with radius 1 centered at the origin. Because of this, and for other reasons that will become apparent later in this book, the balls of radius 1 centered at the origin play a special role in the theory of normed spaces and are therefore given special names.

1.2.3 Definition. Let X be a normed space. The *closed unit ball* of X is $\{ x : x \in X,\ \|x\| \le 1 \}$ and is denoted by B_X. The *open unit ball* of X is $\{ x : x \in X,\ \|x\| < 1 \}$. The *unit sphere* of X is $\{ x : x \in X,\ \|x\| = 1 \}$ and is denoted by S_X.

It is time for some examples.

1.2.4 Example: $\mathbb{F}$. The prototype for all norms is the absolute value function on $\mathbb{F}$. It is the norm implied whenever $\mathbb{F}$ is treated as a normed space without the norm being specified. Notice that this norm induces the standard metric on $\mathbb{F}$.

1.2.5 Example: $\mathbb{F}^n$. If n is a positive integer, then the vector space $\mathbb{F}^n$ of all n-tuples of scalars is a normed space with the *Euclidean norm* given by the formula

$$\|(\alpha_1, \dots, \alpha_n)\| = \left(\sum_{j=1}^{n} |\alpha_j|^2\right)^{1/2}.$$

This normed space is called *Euclidean n-space.* By convention, *Euclidean 0-space* $\mathbb{F}^0$ is a zero-dimensional normed space with its Euclidean norm (and, for that matter, only norm) defined by the formula $\|0\| = 0$. For each nonnegative integer n, the *Euclidean topology* of $\mathbb{F}^n$ is the topology induced by the Euclidean norm. *Henceforth, whenever n is a nonnegative integer and $\mathbb{F}^n$ (or, more specifically, $\mathbb{R}^n$ or $\mathbb{C}^n$) is treated as a normed space without the norm being specified, the Euclidean norm is implied.*

1.2.6 Example: $L_p(\Omega, \Sigma, \mu)$, $1 \le p \le \infty$. Let μ be a positive (that is, nonnegative-extended-real-valued) measure on a σ-algebra Σ of subsets of a set Ω. For each p such that $1 \le p \le \infty$, the Lebesgue space $L_p(\Omega, \Sigma, \mu)$ is a normed space with the norm $\|\cdot\|_p$ given by letting

$$\|f\|_p = \begin{cases} \left(\int_\Omega |f|^p \, d\mu\right)^{1/p} & \text{if } 1 \le p < \infty; \\ \inf\{\, t : t > 0,\ \mu(\{\, x : x \in \Omega,\ |f(x)| > t \,\}) = 0 \,\} & \text{if } p = \infty. \end{cases}$$

For each p, the elements of $L_p(\Omega, \Sigma, \mu)$ are equivalence classes of either real- or complex-valued functions, giving, respectively, real or complex normed spaces. In the rare instances in which it is necessary to be specific about the scalar field, the terms "real $L_p(\Omega, \Sigma, \mu)$" and "complex $L_p(\Omega, \Sigma, \mu)$" are used. A similar convention applies to the spaces defined in the following examples. For many of the Lebesgue spaces appearing in this book, the set Ω is the interval $[0, 1]$, the σ-algebra Σ is the collection of Lebesgue-measurable subsets of $[0, 1]$, and the measure μ is Lebesgue measure on $[0, 1]$, in which case $L_p(\Omega, \Sigma, \mu)$ is abbreviated to $L_p[0, 1]$. Other abbreviations such as $L_p[0, \infty)$ have analogous meanings.

There is not universal agreement on whether the p in the notation for $L_p(\Omega, \Sigma, \mu)$ should be written as a subscript or superscript. Similarly, the roles of subscripts and superscripts are often exchanged in the notations for some of the other normed spaces defined later in this section.

1.2.7 Example: ℓ_∞. The collection of all bounded sequences of scalars is clearly a vector space if the vector space operations are given by letting $(\alpha_j) + (\beta_j) = (\alpha_j + \beta_j)$ and $\alpha \cdot (\alpha_j) = (\alpha \cdot \alpha_j)$. For each element (α_j) of this vector space, let

$$\|(\alpha_j)\|_\infty = \sup\{\, |\alpha_j| : j \in \mathbb{N}\,\}.$$

It is easy to check that $\|\cdot\|_\infty$ is a norm. The resulting normed space is called ℓ_∞ (pronounced "little ell infinity"). A convention used for ℓ_∞ and other spaces whose elements are sequences is that the phrase "a sequence in the space" *always* means a sequence of elements of the space, *never* a single element.

Suppose that μ is the *counting measure* on the collection Σ of all subsets of $\mathbb{N}$, that is, the measure given by letting $\mu(A)$ be the number of elements in A. Then the measurable functions from $\mathbb{N}$ into $\mathbb{F}$ are just the sequences of scalars, and $L_\infty(\mathbb{N}, \Sigma, \mu)$ is ℓ_∞. As with ℓ_∞, the normed spaces of the next two examples can be viewed as Lebesgue spaces by considering the counting measures on the sets $\mathbb{N}$, $\{1, \ldots, n\}$ where $n \in \mathbb{N}$, and $\varnothing$. See Appendix C for an alternative derivation of these spaces that does not use measure theory.

1.2.8 Example: ℓ_p, $1 \le p < \infty$. Let p be a real number such that $p \ge 1$. The collection of all sequences (α_j) of scalars for which $\sum_{j=1}^\infty |\alpha_j|^p$ is finite is a vector space with the vector space operations of the preceding example. Let the norm $\|\cdot\|_p$ be defined on this vector space by the formula

$$\|(\alpha_j)\|_p = \left(\sum_{j=1}^\infty |\alpha_j|^p\right)^{1/p}.$$

The resulting normed space is called ℓ_p (pronounced "little ell p").

1.2.9 Example: ℓ_p^n, $1 \le p \le \infty$. Let p be such that $1 \le p \le \infty$ and let n be a positive integer. Define a norm on the vector space $\mathbb{F}^n$ by letting

$$\|(\alpha_1, \ldots, \alpha_n)\|_p = \begin{cases} \left(\sum_{j=1}^n |\alpha_j|^p\right)^{1/p} & \text{if } 1 \le p < \infty; \\ \max\{|\alpha_1|, \ldots, |\alpha_n|\} & \text{if } p = \infty. \end{cases}$$

The resulting normed space is called ℓ_p^n (pronounced "little ell p n"). Notice that ℓ_2^n is just Euclidean n-space. By convention, the space ℓ_p^0 is Euclidean 0-space. It is also possible to represent ℓ_p^0 as a Lebesgue space by using the fact that there is exactly one scalar-valued function on the empty set, namely, the function represented by $\varnothing \times \mathbb{F}$; see [65, p. 11, Ex. 3]. It follows that ℓ_p^0 can be defined to be $L_p(\varnothing, \{\varnothing\}, \mu)$, where of course $\mu(\varnothing) = 0$.

1.2.10 Example: $C(K)$. Let K be a compact Hausdorff space. Then the collection of all scalar-valued continuous functions on K is a vector space under the operations given by the usual addition of functions and multiplication of functions by scalars. Notice that if K is empty, then this vector space has the single element represented by $K \times \mathbb{F}$; see the reference cited in the preceding example. For each member f of this vector space, let

$$\|f\|_\infty = \begin{cases} \max\{\,|f(x)| : x \in K\,\} & \text{if } K \neq \varnothing; \\ 0 & \text{if } K = \varnothing. \end{cases}$$

It is easy to check that this defines a norm on the vector space. The resulting normed space is denoted by $C(K)$. By analogy with $L_p[0,1]$, the abbreviation $C[0,1]$ represents $C([0,1])$.

1.2.11 Example: $\mathrm{rca}(K)$. Let K be a compact Hausdorff space, and let $\mathrm{rca}(K)$ be the normed space formed from the collection of all regular finite scalar-valued Borel measures on K by defining the vector space operations in the obvious way and by letting $\|\mu\|$ be the total variation of μ on K whenever μ is one of these measures. It follows easily from elementary properties of total variation that $\mathrm{rca}(K)$ really is a normed space; see [67] or [202] for details. Incidentally, the notation $\mathrm{rca}(K)$ comes from the fact that the members of the space are regular countably additive set functions.

It turns out that each of the norms in the preceding examples induces a complete metric. For $\mathbb{F}$, this is a basic fact from analysis; for $C(K)$, this is just a special case of the fact that every uniformly Cauchy sequence of scalar-valued continuous functions on a topological space converges uniformly to a continuous function; and for all of the other spaces except $\mathrm{rca}(K)$, this follows from the completeness of the metrics of Lebesgue spaces. The completeness of $\mathrm{rca}(K)$ is not difficult to prove from basic principles, as is done in Exercise 1.28, and also follows from a result in Section 1.10 about the completeness of dual spaces, as is mentioned in the comments following Theorem 1.10.7. It is also not difficult to prove the completeness of all of the spaces ℓ_p and ℓ_p^n, and therefore of all of the finite-dimensional Euclidean spaces, without invoking the general fact that every Lebesgue space is a complete metric space. See the proof of Theorem C.10 in Appendix C.

It is finally time for the definition of the objects that are the main focus of this book.

1.2.12 Definition. A *Banach norm* or *complete norm* is a norm that induces a complete metric. A normed space is a *Banach space* or *B-space* or *complete normed space* if its norm is a Banach norm.

Each subspace of a normed space X is obviously a normed space with the norm it inherits from X. Just as clearly, each *closed* subspace of a Banach space is itself a Banach space with the inherited norm.

1.2.13 Example: c_0. Let c_0 be the collection of all sequences of scalars that converge to 0, with the same vector space operations and norm as ℓ_∞. Then c_0 is a Banach space since it is a closed subspace of ℓ_∞.

1.2.14 Example: $L_p(\mathbb{T})$ and H_p, $1 \le p \le \infty$. Fix p such that $1 \le p \le \infty$. Let $\mathbb{T}$ be the unit circle $\{ z : z \in \mathbb{C}, |z| = 1 \}$ in the complex plane, and let X be the set of all complex-valued functions f on $\mathbb{T}$ with the property that if $g \colon [-\pi, \pi) \to \mathbb{C}$ is defined by letting $g(t) = f(e^{it})$, then $g \in L_p[-\pi, \pi)$. As would be expected, two functions f_1 and f_2 in X are considered to be the same if $f_1(e^{it}) = f_2(e^{it})$ for almost all t in $[-\pi, \pi)$. Since $L_p[-\pi, \pi)$ is a Banach space, it is clear that X is also a Banach space with the obvious vector space operations and the norm $\|\cdot\|_p$ given by letting

$$\|f\|_p = \begin{cases} \left(\dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} |f(e^{it})|^p \, dt \right)^{1/p} & \text{if } 1 \le p < \infty; \\ \|g\|_\infty \text{ (where } g \text{ is as above)} & \text{if } p = \infty. \end{cases}$$

This Banach space is called $L_p(\mathbb{T})$. Notice that $L_p(\mathbb{T})$ is essentially just $L_p[-\pi, \pi)$, except that $[-\pi, \pi)$ has been identified with $\mathbb{T}$ and Lebesgue measure λ has been replaced by *normalized Lebesgue measure* $(2\pi)^{-1}\lambda$ so that the measure of $\mathbb{T}$ is 1.

Suppose that $f \in L_p(\mathbb{T})$. For each integer n, the n^{th} *Fourier coefficient* $\hat{f}(n)$ of f is defined by the formula

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{it}) e^{-int} \, dt.$$

If n is an integer and (f_j) is a sequence in $L_p(\mathbb{T})$ that converges to some f in $L_p(\mathbb{T})$, then $|\hat{f}_j(n) - \hat{f}(n)| \le \|f_j - f\|_1 \le \|f_j - f\|_p \to 0$ as $j \to \infty$, and so $\lim_j \hat{f}_j(n) = \hat{f}(n)$. It follows that

$$\{ f : f \in L_p(\mathbb{T}),\ \hat{f}(n) = 0 \text{ whenever } n < 0 \}$$

is a closed subspace of $L_p(\mathbb{T})$ and hence is itself a Banach space with the norm inherited from $L_p(\mathbb{T})$. This Banach space is called the *Hardy space* H_p.

Actually, the space H_p is usually defined to be the class of all functions analytic in the open unit disc of $\mathbb{C}$ and satisfying a certain condition, depending on p, that restricts their growth near the boundary of that disc. It turns out that for each such function f there is a *boundary function* $f_{\mathbb{T}} \colon \mathbb{T} \to \mathbb{C}$ such that $\lim_{r \to 1^-} f(re^{it}) = f_{\mathbb{T}}(e^{it})$ for almost all t in $[-\pi, \pi)$, and that the resulting collection of boundary functions is a subspace of $L_p(\mathbb{T})$. It is common for H_p to be viewed as a subspace of $L_p(\mathbb{T})$ by identifying each function satisfying the growth restriction with its boundary function. That is how the definition given above of H_p as a subspace of $L_p(\mathbb{T})$ is obtained. See [68] or [202] for details.

The following vector space, with various norms, is useful for generating counterexamples.

1.2.15 Example. The *vector space of finitely nonzero sequences* is the collection of all sequences of scalars that have only finitely many nonzero terms, with the obvious vector space operations. This vector space is clearly a subspace of c_0 and of each space ℓ_p such that $1 \le p \le \infty$, and is, in fact, dense in all of these except ℓ_∞. See Exercise 1.24.

1.2.16 Definition. Let X be a vector space of sequences of scalars that has the vector space of finitely nonzero sequences as a subspace. For each positive integer n, the n^{th} *standard coordinate vector* e_n of X is the element of X that has 1 as its n^{th} term and all other terms 0. If X is the vector space underlying a normed space and each of its standard coordinate vectors has norm 1, then each standard coordinate vector is also called a *standard unit vector* of X.

Since the collection of standard coordinate vectors is a basis for the vector space of finitely nonzero sequences, it follows that every vector space that has the vector space of finitely nonzero sequences as a subspace is infinite-dimensional. In particular, the Banach spaces c_0 and ℓ_p such that $1 \le p \le \infty$ are all infinite-dimensional. Similarly, the Banach spaces $C[0,1]$ and $L_p[0,1]$ such that $1 \le p \le \infty$ are infinite-dimensional since each has in it the linearly independent sequence (x_n), where $x_n(t) = t^n$ whenever $n \in \mathbb{N}$ and $0 \le t \le 1$.

Exercises

1.18 With all notation as in the first paragraph of this section, show that if $(x_1, x_2] = (y_1, y_2]$, then $x_1 = y_1$ and $x_2 = y_2$. (This is easy when $x_1 = x_2$. Suppose instead that $x_1 \ne x_2$. It may help to show first that $[x_1, x_2] = [y_1, y_2]$, where $[x_1, x_2] = \{\, y : y = (1-t)x_1 + tx_2,\ 0 \le t \le 1 \,\}$. To this end, notice that $(x_1, x_2]$ and $[x_1, x_2]$ are both convex and that $[x_1, x_2]$ has exactly one more point than does $(x_1, x_2]$. In how many ways can $(x_1, x_2]$ be augmented by one point so that the resulting set is convex?)

1.19 The definition of a norm contains some redundancies. Prove that an equivalent definition is obtained by replacing (1) in that definition with

(1′) $\|x\| \ne 0$ whenever $x \ne 0$.

1.20 A *sphere* in a metric space X is a set of the form $\{\, y : y \in X,\ d(x, y) = r \,\}$, where $x \in X$ and $r > 0$.

(a) Prove that if X is a normed space, then every closed ball is the closure of the corresponding open ball, every open ball is the interior of the corresponding closed ball, and every sphere is the boundary of the corresponding open and closed balls.

(b) Find a metric space X such that the three conclusions of part (a) all fail to hold for X.

1.21 Let X be a normed space. Prove that if (B_n) is a sequence of balls in X such that $B_n \supseteq B_{n+1}$ for each n, then the centers of the balls form a Cauchy sequence. Give an example to show that this result can fail if X is only assumed to be a metric space.

1.22 The purpose of this exercise is to show that the Banach spaces of Examples 1.2.7, 1.2.8, and 1.2.9 can be obtained by subspace arguments beginning with the Banach spaces $L_p[0, \infty)$.

(a) Show that ℓ_∞ can be identified with a closed subspace of $L_\infty[0, \infty)$ by identifying (α_j) with the function whose value on each interval $[j-1, j)$ is α_j. Conclude that ℓ_∞ is a Banach space.

(b) Use a similar argument to prove that ℓ_p is a Banach space when $1 \leq p < \infty$.

(c) For each p such that $1 \leq p \leq \infty$ and each nonnegative integer n, identify ℓ_p^n with a closed subspace of ℓ_p and conclude that ℓ_p^n is a Banach space.

(d) Give analogous arguments for (a) and (b) based on the spaces $L_p[0, 1]$ instead of $L_p[0, \infty)$.

1.23 Let X be the vector space of all continuous functions from $[0, 1]$ into $\mathbb{F}$. For each p such that $1 \leq p \leq \infty$, let $\|\cdot\|_p$ be the norm that X inherits from $L_p[0, 1]$. For which values of p is this a Banach norm?

1.24 Let X be the vector space of finitely nonzero sequences. Show that X is a dense subspace of c_0 and of ℓ_p when $1 \leq p < \infty$, but not of ℓ_∞.

1.25 Let c be the collection of all convergent sequences of scalars with the vector space operations and norm as given for ℓ_∞. Show that c is a Banach space.

1.26 (a) In each of the spaces c_0 and ℓ_p such that $1 \leq p \leq \infty$, identify the linear hull $\langle \{ e_n : n \in \mathbb{N} \} \rangle$ of the collection of unit vectors.

(b) Do the same for $[\{ e_n : n \in \mathbb{N} \}]$.

(c) Do the same for $\mathrm{co}(\{ e_n : n \in \mathbb{N} \})$.

1.27 (a) Suppose that the unit sphere of ℓ_1 contains the sequence $(x^{(n)})$ and the element x. Show that $(x^{(n)})$ converges to x if and only if $\lim_n x_j^{(n)} = x_j$ for each positive integer j.

(b) Show that (a) fails if the requirement that $\|x\| = 1$ is removed.

(c) Show that (a) fails if ℓ_1 is replaced by c_0.

1.28 Suppose that Σ is a σ-algebra of subsets of a set Ω. Let $\mathrm{ca}(\Omega, \Sigma)$ be the normed space formed from the collection of all finite scalar-valued measures on Σ by using the obvious vector space operations and by letting the norm of a measure be its total variation. It is not difficult to show that $\mathrm{ca}(\Omega, \Sigma)$ really is a normed space; see [202, pp. 116–119] for the details.

(a) Prove that $\mathrm{ca}(\Omega, \Sigma)$ is a Banach space.

(b) Now suppose that $\mathfrak{B}$ is the σ-algebra of Borel subsets of a compact Hausdorff space K. Prove that $\mathrm{rca}(K)$ is a closed subspace of $\mathrm{ca}(K, \mathfrak{B})$ and is therefore a Banach space.

1.29 Fix a positive real number p. A scalar-valued function f on $[0, 1]$ is said to satisfy a *Lipschitz condition of order* p if there is a nonnegative real number M such that $|f(s) - f(t)| \leq M|s - t|^p$ whenever $s, t \in [0, 1]$. The collection of all such functions is denoted by $\mathrm{Lip}\, p$.

(a) Show that $\mathrm{Lip}\, p$ is a subspace of the vector space of continuous scalar-valued functions on $[0, 1]$. Show that if $p > 1$, then $\mathrm{Lip}\, p$ consists only of the constant scalar-valued functions on $[0, 1]$.

(b) Define a function $\|\cdot\|\colon \mathrm{Lip}\, p \to \mathbb{R}$ by the formula

$$\|f\| = |f(0)| + \sup\left\{ \frac{|f(s) - f(t)|}{|s - t|^p} : s, t \in [0, 1],\ s \neq t \right\}.$$

Prove that $\|\cdot\|$ is a norm. For the rest of this exercise, $\mathrm{Lip}\, p$ is to be treated as a normed space having this norm.

(c) Prove that $\mathrm{Lip}\, p$ is a Banach space.

(d) Let $\mathrm{lip}\, p$ be the collection of all f in $\mathrm{Lip}\, p$ such that

$$\lim_{\delta \to 0^+} \sup\left\{ \frac{|f(s) - f(t)|}{|s - t|^p} : s, t \in [0, 1],\ s \neq t,\ |s - t| \leq \delta \right\} = 0.$$

Prove that $\mathrm{lip}\, p$ is a Banach space under the norm inherited from $\mathrm{Lip}\, p$.

1.3 First Properties of Normed Spaces

The first order of business is to obtain a few basic continuity results.

1.3.1 Proposition. *Let X be a normed space. Then $\big|\|x\| - \|y\|\big| \leq \|x - y\|$ whenever $x, y \in X$. Thus, the function $x \mapsto \|x\|$ is continuous from X into $\mathbb{R}$.*

PROOF. If $x, y \in X$, then $\|x\| \leq \|x - y\| + \|y\|$ and $\|y\| \leq \|y - x\| + \|x\| = \|x - y\| + \|x\|$, so $\big|\|x\| - \|y\|\big| = \max\{\, \|x\| - \|y\|,\ \|y\| - \|x\|\, \} \leq \|x - y\|$. ∎

1.3.2 Proposition. *Let X be a normed space.*

(a) *Addition of vectors is a continuous operation from $X \times X$ into X.*

(b) *Multiplication of vectors by scalars is a continuous operation from $\mathbb{F} \times X$ into X.*

PROOF. Suppose that $\alpha, \alpha_0 \in \mathbb{F}$, that $x, y, x_0, y_0 \in X$, that $\epsilon > 0$, and that $|\alpha - \alpha_0| < \epsilon$, $\|x - x_0\| < \epsilon$, and $\|y - y_0\| < \epsilon$. Then

$$\|(x + y) - (x_0 + y_0)\| \leq \|x - x_0\| + \|y - y_0\| < 2\epsilon$$

and

$$\begin{aligned}\|\alpha x - \alpha_0 x_0\| &\le \|\alpha x - \alpha_0 x\| + \|\alpha_0 x - \alpha_0 x_0\| \\ &= |\alpha - \alpha_0|\,\|x\| + |\alpha_0|\,\|x - x_0\| \\ &< \epsilon(\|x_0\| + \epsilon) + |\alpha_0|\epsilon,\end{aligned}$$

from which (a) and (b) follow. ■

1.3.3 Corollary. *Let x_0 be an element of a normed space X and let α_0 be a nonzero scalar. Then the maps $x \mapsto x + x_0$ and $x \mapsto \alpha_0 x$ are homeomorphisms from X onto itself. Consequently, if A is a subset of X that is open, or closed, or compact, then $x_0 + A$ and $\alpha_0 A$ also have that property. If A and U are subsets of X and U is open, then $A + U$ is open.*

PROOF. The continuity of the vector space operations of X implies that of the maps $x \mapsto x + x_0$ and $x \mapsto \alpha_0 x$, as well as that of the respective inverse maps $x \mapsto x + (-x_0)$ and $x \mapsto \alpha_0^{-1}x$. The remark about translates and nonzero scalar multiples of sets that are open, closed, or compact now follows, since homeomorphisms preserve these properties. Finally, if A and U are subsets of X and U is open, then $A + U = \bigcup\{\, a + U : a \in A\,\}$, a union of open sets. ■

1.3.4 Corollary. *Let S be a topological space and X a normed space. Then the collection of all continuous functions from S into X is a vector space over $\mathbb{F}$ if sums and scalar multiples of functions are defined in the usual way: $(f+g)(s) = f(s) + g(s)$ and $(\alpha \cdot f)(s) = \alpha \cdot f(s)$ for each s in S.*

PROOF. If α is a scalar and f and g are continuous functions from S into X, then the continuity of the vector space operations of X implies that the maps $s \mapsto \big(f(s), g(s)\big) \mapsto (f+g)(s)$ and $s \mapsto f(s) \mapsto \alpha f(s)$ from S into X are both continuous, that is, that $f + g$ and αf are continuous. The verification of the vector space axioms is then easy. Notice that the zero of this vector space is the map $s \mapsto 0$ and that $(-f)(s) = -\big(f(s)\big)$ for each f in this vector space and each s in S. ■

Because of the continuity of the vector space operations, it is possible to develop an interesting theory of series in normed spaces.

1.3.5 Definition. Suppose that (x_n) is a sequence in a normed space. Then the *series generated by* (x_n) is the sequence $\big(\sum_{n=1}^m x_n\big)_{m=1}^\infty$. For each positive integer m, the m^{th} term $\sum_{n=1}^m x_n$ of this sequence of sums is the m^{th} *partial sum* of the series. If the series converges, that is, if $\lim_m \sum_{n=1}^m x_n$ exists, then this limit is the *sum* of the series and is denoted by $\sum_{n=1}^\infty x_n$ or by $\sum_n x_n$.

1.3.6 Example. Suppose that X is ℓ_p, where $1 \le p < \infty$, or c_0. Let (e_n) be the sequence of standard unit vectors in X. It is easy to check that

$(\alpha_n) = \sum_n \alpha_n e_n$ whenever $(\alpha_n) \in X$. This result does not extend to ℓ_∞, for if (α_n) is a member of ℓ_∞ whose terms do not tend to 0, then the sequence $\left(\sum_{n=1}^m \alpha_n e_n\right)_{m=1}^\infty$ is not Cauchy and therefore cannot converge.

If a sequence (x_n) in a normed space generates a convergent series, then it is often said that $\sum_n x_n$ is a convergent series or that $\sum_n x_n$ converges, even though $\sum_n x_n$ is only the limit of the series and is not the series itself. The term *infinite sum* is also used for both a convergent series and its limit. Even when the series generated by the sequence (x_n) might not converge, the sequence of partial sums is often called the *formal series* $\sum_n x_n$. To complicate things just a bit more, the notation $\sum_n x_n$ is also used occasionally for the sum of a finite list $x_1, \dots, x_m$ of objects. In any case, the context should always prevent confusion.

1.3.7 Proposition. *Let X and Y be normed spaces.*

(a) *If $\sum_n x_n$ converges in X, then $x_n \to 0$.*
(b) *If $\sum_n x_n$ and $\sum_n y_n$ both converge in X, then so does $\sum_n(x_n+y_n)$, and $\sum_n(x_n + y_n) = \sum_n x_n + \sum_n y_n$.*
(c) *If $\sum_n x_n$ converges in X and α is a scalar, then $\sum_n \alpha x_n$ converges, and $\sum_n \alpha x_n = \alpha \sum_n x_n$.*
(d) *If $\sum_n x_n$ converges in X and T is a continuous linear operator from X into Y, then $\sum_n Tx_n$ converges in Y, and $\sum_n Tx_n = T\left(\sum_n x_n\right)$.*
(e) *If $\sum_n x_n$ is a finite or infinite sum in X, then $\|\sum_n x_n\| \le \sum_n \|x_n\|$.*

PROOF. Part (a) follows immediately from the fact that if $\sum_n x_n$ is a convergent series in X, then each x_n besides the first is the difference of two consecutive terms of the Cauchy sequence of partial sums of the series. Notice next that if α is a scalar and both $\sum_n x_n$ and $\sum_n y_n$ converge in X, then the continuity of the vector space operations implies that

$$\sum_{n=1}^m (\alpha x_n + y_n) = \alpha \sum_{n=1}^m x_n + \sum_{n=1}^m y_n \to \alpha \sum_n x_n + \sum_n y_n$$

as $m \to \infty$, giving (b) and (c). A similar argument, with α replaced by T and references to y_n omitted, proves (d). For (e), repeated applications of the triangle inequality show that $\|\sum_{n=1}^m x_n\| \le \sum_{n=1}^m \|x_n\|$ for each finite sum $\sum_{n=1}^m x_n$ in X. If $\sum_n x_n$ is an infinite sum in X, then the continuity of the map $x \mapsto \|x\|$ assures that $\|\sum_{n=1}^m x_n\| \to \|\sum_n x_n\|$ as $m \to \infty$, so the desired inequality for infinite sums can be obtained by letting m tend to infinity in the inequality $\|\sum_{n=1}^m x_n\| \le \sum_{n=1}^m \|x_n\|$. ■

Of course, the converse of (a) is not in general true since the harmonic series diverges. It should also be noted that the sum $\sum_n \|x_n\|$ in (e) does not have to be finite when $\sum_n x_n$ is a convergent infinite sum; otherwise, there would be no need to develop a theory of absolutely and conditionally convergent series in $\mathbb{F}$. A similar theory exists for arbitrary normed spaces, and has as one of its basic results a very nice characterization of Banach spaces among all normed spaces.

1.3.8 Definition. Let $\sum_n x_n$ be a formal series in a normed space.

(a) The series $\sum_n x_n$ is *absolutely convergent* if $\sum_n \|x_n\|$ converges.

(b) The series $\sum_n x_n$ is *unconditionally convergent* if $\sum_n x_{\pi(n)}$ converges for each permutation π of $\mathbb{N}$.

(c) The series $\sum_n x_n$ is *conditionally convergent* if it is convergent but not unconditionally convergent.

The notions of absolute and unconditional convergence are equivalent in $\mathbb{F}$; see [140]. In fact, it was proved by A. Dvoretzky and C. A. Rogers [70] in a 1950 paper that the equivalence of absolute and unconditional convergence of series actually characterizes finite-dimensional normed spaces among all normed spaces.

1.3.9 Theorem. *A normed space X is a Banach space if and only if each absolutely convergent series in X converges.*

PROOF. Suppose that X is not a Banach space. Let (x_n) be a nonconvergent Cauchy sequence in X. For each positive integer j, there is a positive integer n_j such that $\|x_n - x_m\| \leq 2^{-j}$ if $n, m \geq n_j$. It can be assumed that $n_{j+1} > n_j$ for each j. Since a limit for a subsequence of a Cauchy sequence must be a limit for the entire sequence, the subsequence (x_{n_j}) of (x_n) has no limit. Therefore, the series $\sum_j (x_{n_{j+1}} - x_{n_j})$ is not convergent, since $\sum_{j=1}^k (x_{n_{j+1}} - x_{n_j}) = x_{n_{k+1}} - x_{n_1}$ for each positive integer k. However, this series is absolutely convergent, since $\sum_j \|x_{n_{j+1}} - x_{n_j}\| \leq \sum_j 2^{-j} = 1$.

Conversely, suppose that X is a Banach space and that $\sum_n x_n$ is an absolutely convergent series in X. If $m_1, m_2 \in \mathbb{N}$ and $m_2 > m_1$, then

$$\left\| \sum_{n=1}^{m_2} x_n - \sum_{n=1}^{m_1} x_n \right\| \leq \sum_{n=m_1+1}^{m_2} \|x_n\| = \sum_{n=1}^{m_2} \|x_n\| - \sum_{n=1}^{m_1} \|x_n\|,$$

from which it follows that the partial sums of $\sum_n x_n$ form a Cauchy sequence and therefore that $\sum_n x_n$ converges. ■

See Exercise 1.31 for some further observations about absolute and unconditional convergence of series in Banach spaces. The reader interested in unconditional convergence should also see Propositions 4.2.1 and 4.2.3. While the proofs of those two propositions rely only on material that has already been covered, the results have been postponed until Section 4.2 since they are not needed before then.

If X is a vector space with a topology and $A \subseteq X$, then it need not be true that $[A] = \overline{\langle A \rangle}$ or that $\overline{\text{co}}(A) = \overline{\text{co}(A)}$; see Exercise 1.12. It is an easy consequence of the next theorem that these equalities do hold if X is a normed space. See also Exercise 1.33.

1.3.10 Theorem. *Let X be a normed space.*

(a) *If S is a subspace of X, then $\overline{S}$ is a subspace of X.*

(b) *If C is a convex subset of X, then both $\overline{C}$ and C° are convex.*

PROOF. Let S be a subspace of X. Suppose that $x, y \in \overline{S}$ and that α is a scalar. Let (x_n) and (y_n) be sequences in S converging to x and y respectively. Then $(\alpha x_n + y_n)$ is a sequence in S converging to $\alpha x + y$, and so $\alpha x + y \in \overline{S}$. Since $\overline{S}$ is nonempty and contains the sum of each pair of its elements as well as each scalar multiple of each of its elements, it follows that $\overline{S}$ is a subspace of X.

Suppose that C is a convex subset of X and that $0 < t < 1$. If $x, y \in \overline{C}$, then an argument similar to that just given shows that $tx + (1-t)y \in \overline{C}$, so $\overline{C}$ is convex. To see that C° is convex, notice first that $tC^\circ + (1-t)C^\circ$ is open by Corollary 1.3.3. Since $tC^\circ + (1-t)C^\circ \subseteq tC + (1-t)C = C$, it follows that $tC^\circ + (1-t)C^\circ \subseteq C^\circ$, so C° is convex. ■

1.3.11 Corollary. *Let A be a subset of a normed space. Then $[A] = \overline{\langle A \rangle}$ and $\overline{\text{co}}(A) = \overline{\text{co}(A)}$.*

PROOF. Let X be the normed space. It follows from the theorem that $\overline{\langle A \rangle}$ is a closed subspace of X that includes A, so $\overline{\langle A \rangle} \supseteq [A]$. Since $[A]$ is a subspace of X that includes A, it is also true that $\langle A \rangle \subseteq [A]$. Therefore $\overline{\langle A \rangle} \subseteq [A]$ since $[A]$ is closed, and so $\overline{\langle A \rangle} = [A]$. The other equality is proved similarly. ■

It is not in general true that balls in a vector space with a metric must be convex; see Exercise 1.35. However, balls in a normed space always have this property.

1.3.12 Proposition. *Every ball in a normed space, whether open or closed, is convex.*

PROOF. Suppose that B is a ball in a normed space and that x and r are, respectively, the center and radius of B. If $y, z \in B$ and $0 < t < 1$, then

$$\begin{aligned} \|ty + (1-t)z - x\| &= \|ty + (1-t)z - tx - (1-t)x\| \\ &\le t\|y - x\| + (1-t)\|z - x\| \\ &\le r, \end{aligned}$$

where the last inequality is strict if B is open. Therefore $ty + (1-t)z \in B$, and so B is convex. ■

Balls centered at the origin of a normed space have some additional useful features.

1.3.13 Proposition. *Every ball centered at the origin of a normed space, whether open or closed, is balanced and absorbing.*

PROOF. Suppose that B is an open or closed ball centered at the origin of a normed space X and that r is the radius of B. It is clear that B is balanced. If $x \in X$ and $t > r^{-1}\|x\|$, then $\|t^{-1}x\| = t^{-1}\|x\| < r$, so $t^{-1}x \in B$ and therefore $x \in tB$. Thus, the set B is absorbing. ■

One consequence of the preceding two propositions is that every neighborhood of the origin in a normed space includes a closed, convex, absorbing set since it includes a closed ball centered at the origin. The following partial converse of this is a special version of a result called the Baire category theorem, a more general form of which is given in Section 1.5. Notice that this theorem is stated only for Banach spaces, not arbitrary normed spaces. As is shown in Exercise 1.36, there are incomplete normed spaces having closed, convex, absorbing subsets with empty interiors.

1.3.14 Theorem. *Every closed, convex, absorbing subset of a Banach space includes a neighborhood of the origin.*

PROOF. Let C be a closed, convex, absorbing subset of a Banach space, and let $D = C \cap (-C)$. It is enough to show that D includes a neighborhood of the origin. If A is a nonempty subset of D, then

$$0 \in \frac{1}{2}A + \frac{1}{2}(-A) \subseteq \frac{1}{2}D + \frac{1}{2}(-D) = \frac{1}{2}D + \frac{1}{2}D = D,$$

where the last equality comes from the convexity of D. Thus, it is enough to prove that $D^\circ \neq \varnothing$, for then the neighborhood $\frac{1}{2}D^\circ + \frac{1}{2}(-D^\circ)$ of the origin must be included in D.

Suppose that D has empty interior. For each positive integer n, the set nD is closed and has empty interior, and so $X \setminus nD$ is an open set that is dense in X. Let B_1 be a closed ball in $X \setminus D$ with radius no more than 1. Since $(X \setminus 2D) \cap B_1^\circ$ is a nonempty open set, there is a closed ball B_2 in $B_1 \setminus 2D$ with radius no more than $1/2$. There is a closed ball B_3 in $B_2 \setminus 3D$ with radius no more than $1/3$. Continuing in the obvious way yields a sequence (B_n) of closed balls such that these all hold for each positive integer n: $B_n \cap nD = \varnothing$, the radius of B_n is no more than $1/n$, and $B_n \supseteq B_m$ if $n \leq m$. It follows that the centers of the balls form a Cauchy sequence whose limit x is in each of the balls and hence is in $X \setminus nD$ for each n. However, the set C is absorbing, so there is a positive real number s such that if $t > s$, then $x, -x \in tC$ and therefore $x \in tD$. It follows that $x \in nD$ for some positive integer n, a contradiction that proves the theorem. ■

In the preceding proof, it is not really necessary to force the radii of the balls to decrease to 0. The fact that $B_n \supseteq B_m$ whenever $n \le m$ is enough to assure that their centers form a Cauchy sequence. See Exercise 1.21.

Exercises

1.30 Let K be a compact Hausdorff space and let X be a normed space. By Corollary 1.3.4, the collection of all continuous functions from K into X is a vector space when functions are added and multiplied by scalars in the usual way. Define a norm on this vector space by the formula

$$\|f\|_\infty = \begin{cases} \max\{\, \|f(x)\| : x \in K \,\} & \text{if } K \neq \varnothing; \\ 0 & \text{if } K = \varnothing. \end{cases}$$

The resulting normed space is denoted by $C(K, X)$.

(a) Show that $\|\cdot\|_\infty$ is in fact a norm on $C(K, X)$.

(b) Show that if X is a Banach space, then so is $C(K, X)$.

1.31 Let $\sum_n x_n$ be a formal series in a Banach space.

(a) Prove that if the series is absolutely convergent, then it is unconditionally convergent.

(b) Give an example to show that the converse of (a) need not hold. (The space c_0 is a good place to look.)

1.32 Identify $\overline{\text{co}}(\{\, e_n : n \in \mathbb{N} \,\})$ in ℓ_1. Exercises 1.26 (c) and 1.27 (a) may help.

1.33 This exercise shows that the order in which closures and hulls are taken in Corollary 1.3.11 is important.

(a) Prove that if A is a subset of a normed space, then $\overline{\text{co}}(A) \supseteq \text{co}(\overline{A})$ and $[A] \supseteq \langle \overline{A} \rangle$.

(b) Show that if A is the collection of unit vectors in c_0, then the inclusions in (a) are proper.

1.34 Suppose that S is a subspace of a normed space X. It is clear that S° need not be a subspace of X because S° might be empty; consider $\{0\}^\circ$ in $\mathbb{F}$. It is true, however, that S° is a subspace of X when $S^\circ \neq \varnothing$. In fact, a much stronger statement can then be made about S°. What is that statement?

1.35 Euclidean-space intuition might make it tempting to think that balls in a vector space with a metric must be convex. To see that this is not so, let X be the vector space $\mathbb{R}^2$ with the metric given by the formula

$$d\big((\alpha_1, \beta_1), (\alpha_2, \beta_2)\big) = |\alpha_1 - \alpha_2|^{1/2} + |\beta_1 - \beta_2|^{1/2}.$$

Show that d really is a metric on $\mathbb{R}^2$. Let B_X be the closed unit ball of X, that is, the set of all members of X no more than 1 unit from $(0, 0)$. Sketch B_X. Notice that $(1, 0)$ and $(0, 1)$ are in B_X, but $(\frac{1}{2}, \frac{1}{2})$ is not. Thus, the ball B_X is not convex.

1.36 Suppose that X is the vector space underlying ℓ_1, but equipped with the ℓ_∞ norm. Show that $\{(\alpha_n) : (\alpha_n) \in X, \sum_n |\alpha_n| \leq 1\}$ is a closed, convex, absorbing subset of X whose interior is empty. (Notice that, by Theorem 1.3.14, the normed space X cannot be a Banach space.)

1.37 Suppose that a vector space X has a nonempty subset B that is convex, balanced, and has this strong absorbing property: For every nonzero x in X, there is a positive s_x such that $x \in tB$ if $t \geq s_x$ and $x \notin tB$ if $0 \leq t < s_x$. Show that there is a norm $\|\cdot\|_B$ on X for which B is the closed unit ball.

1.38 Suppose that T is a linear operator from a Banach space X into a normed space Y. Show that if $T^{-1}(B_Y)$ is closed, then T is continuous at 0. (It can then be concluded from Theorem 1.4.2 that T is actually continuous everywhere.)

1.39 The *core* of a subset A of a vector space X is the collection of all points y in A with this property: For each x in X, there is a positive $\delta_{x,y}$ such that $y + tx \in A$ whenever $0 \leq t < \delta_{x,y}$. Prove that if X is a Banach space, then the core of each closed convex subset of X is the interior of that set.

1.4 Linear Operators Between Normed Spaces

Linear operators between vector spaces were discussed briefly in Section 1.1. The purpose of this section is to derive a few basic properties of linear operators between normed spaces, with special emphasis on the continuous ones.

Recall that a subset of a metric space is *bounded* if it is either empty or included in some ball. (Making a special case of the empty set is necessary only when the metric space is itself empty.) Each compact subset K of a metric space is bounded, since the metric space is either empty or the union of an increasing sequence of open balls which thus form an open covering for K, forcing K to lie inside one of the balls. Also, every Cauchy sequence in a metric space is bounded, since the sequence from some term onward lies inside a ball of radius 1, and increasing the radius of that ball, if necessary, causes it to engulf the rest of the terms.

It is easy to check that a subset A of a normed space is bounded if and only if there is a nonnegative number M such that $\|x\| \leq M$ for each x in A.

1.4.1 Definition. Let X and Y be normed spaces. A linear operator T from X into Y is *bounded* if $T(B)$ is a bounded subset of Y whenever B is a bounded subset of X. The collection of all bounded linear operators from X into Y is denoted by $B(X, Y)$, or by just $B(X)$ if $X = Y$.

A function into a metric space is often called bounded when its range is a bounded set. It is quite possible for a linear operator between normed spaces to be unbounded in this sense and yet bounded in the sense of Definition 1.4.1, an example being the operator $I\colon \mathbb{R} \to \mathbb{R}$ that maps each real number to itself. In fact, if X and Y are normed spaces, then the only member of $L(X,Y)$ with a bounded range is the operator with range $\{0\}$, for the range of each other member of $L(X,Y)$ contains a nonzero vector and all scalar multiples of that vector. Thus, the study of bounded-range linear operators between normed spaces does not lead to a particularly rich theory, and so the term *bounded* when applied to linear operators between normed spaces always has the meaning given in Definition 1.4.1.

As was suggested earlier, the main emphasis in this section is on *continuous* linear operators between normed spaces. The following theorem makes clear the reason for introducing boundedness into this discussion.

1.4.2 Theorem. *Let X and Y be normed spaces and let $T\colon X \to Y$ be a linear operator. Then the following are equivalent.*

(a) *The operator T is continuous.*

(b) *The operator T is continuous at* 0.

(c) *The operator T is uniformly continuous on X.*

(d) *The operator T is bounded.*

(e) *For some neighborhood U of* 0 *in X, the set $T(U)$ is bounded in Y.*

(f) *There is a nonnegative real number M such that $\|Tx\| \le M\|x\|$ for each x in X.*

(g) *The quantity* $\sup\{\, \|Tx\| : x \in B_X \,\}$ *is finite.*

PROOF. Suppose that T is continuous at 0. If $\epsilon > 0$, then there is a positive δ such that $\|Tx\| = \|Tx - T0\| < \epsilon$ whenever $x \in X$ and $\|x\| = \|x - 0\| < \delta$, so $\|Tx_1 - Tx_2\| = \|T(x_1 - x_2)\| < \epsilon$ whenever $x_1, x_2 \in X$ and $\|x_1 - x_2\| < \delta$. Therefore T is uniformly continuous on X, which proves that (b) $\Rightarrow$ (c). It is clear that (c) $\Rightarrow$ (a) $\Rightarrow$ (b), and so (a), (b), and (c) are equivalent.

If (b) holds, then there is an open ball V centered at 0 such that $\|Tx\| < 1$ whenever $x \in V$. For each bounded subset B of X, there is a positive t_B such that $B \subseteq t_B V$, and so $\|Tx\| < t_B$ if $x \in B$. Thus, the operator T is bounded, and so (b) $\Rightarrow$ (d). If (d) holds, then so does (e), as can be seen by letting U be the open unit ball of X, a bounded set. Suppose that (e) holds. Let r be a positive number small enough that the closed ball B_r of radius r and center 0 is included in U, and let $M_0 = \sup\{\, \|Tx\| : x \in B_r \,\}$. If x is a nonzero member of X, then $r\|x\|^{-1}x \in B_r$, so $\left\|T(r\|x\|^{-1}x)\right\| \le M_0$ and therefore $\|Tx\| \le r^{-1}M_0\|x\|$. This last inequality also holds if $x = 0$, so (e) $\Rightarrow$ (f). It is clear that (f) $\Rightarrow$ (g). Finally, suppose that (g) holds. Let $M = \sup\{\, \|Tx\| : x \in B_X \,\}$. If $\epsilon > 0$ and $x \in \epsilon B_X$, then $\|Tx - T0\| = \|Tx\| \le \epsilon M$, which establishes the continuity of T at 0. Therefore (g) $\Rightarrow$ (b), and so (b), (d), (e), (f), and (g) are equivalent. ■

It follows from the preceding theorem that the collection $B(X, Y)$ of all bounded linear operators from a normed space X into a normed space Y is the same as the collection of all continuous linear operators from X into Y. Convention dictates that a member of $B(X, Y)$ be called bounded rather than continuous, which also explains why the notation $B(X, Y)$ is used for this collection rather than something like $CL(X, Y)$. While $L(X, Y)$ is used in this book to denote the vector space of all linear operators from X into Y, be warned that many authors use $L(X, Y)$ to denote the corresponding space of bounded linear operators.

Let V be the vector space of all continuous functions from a normed space X into a normed space Y with addition of functions and multiplication of functions by scalars defined in the usual way; see Corollary 1.3.4. Then $B(X, Y)$ is clearly a subspace of V, and so is itself a vector space. It would be nice if $B(X, Y)$ could somehow be made into a normed space. To do that, some notion for the "length" or "size" of a member T of $B(X, Y)$ is needed. The quantity $\sup\{ \|Tx\| : x \in B_X \}$ mentioned in part (g) of Theorem 1.4.2 is in a sense a measure of the size of $T(B_X)$, and therefore would seem to be a plausible candidate.

1.4.3 Definition. Let X and Y be normed spaces. For each T in $B(X, Y)$, the *norm* or *operator norm* $\|T\|$ of T is the nonnegative real number $\sup\{ \|Tx\| : x \in B_X \}$. The *operator norm* on $B(X, Y)$ is the map $T \mapsto \|T\|$.

1.4.4 Example. Let X and Y be normed spaces. The *zero operator* from X into Y is the zero element of the vector space $B(X, Y)$, that is, the operator that maps each x in X to the zero element of Y. This operator clearly has norm 0. The *identity operator* on X is the member I of $B(X)$ defined by the formula $Ix = x$. Notice that $\|I\| = 1$ as long as $X \neq \{0\}$. More generally, for each scalar α the linear operator T_α on X given by the formula $T_\alpha x = \alpha x$ is bounded, and has norm $|\alpha|$ if $X \neq \{0\}$. Such operators are called *scalar operators.*

1.4.5 Example. Let X be the vector space of finitely nonzero sequences with the ℓ_∞ norm. Define a linear operator $T\colon X \to \mathbb{F}$ by the formula $T(\alpha_n) = \sum_n \alpha_n$. For each positive integer n, let $x_n = \sum_{j=1}^{n} e_j$. Then $x_n \in B_X$ and $Tx_n = n$ for each n. It follows that the set $T(B_X)$ is not bounded, so neither is the operator T.

1.4.6 Example. Define a linear operator $T\colon C[0,1] \to C[0,1]$ by mapping each f in $C[0,1]$ to its indefinite integral F given by the formula $F(t) = \int_0^t f(s)\,ds$. Then T is bounded, and $\|T\| = 1$. See Exercise 1.40.

Here are some useful characterizations of the operator norm that are sometimes used for its definition.

1.4.7 Proposition. *Suppose that X and Y are normed spaces and that T is a bounded linear operator from X into Y.*

(a) $\|T\| = \sup\{\, \|Tx\| : x \in X,\ \|x\| < 1\,\}$.

(b) *If $X \neq \{0\}$, then* $\|T\| = \sup\{\, \|Tx\| : x \in S_X \,\}$.

(c) *If $x \in X$, then $\|Tx\| \le \|T\|\,\|x\|$. Furthermore, the number $\|T\|$ is the smallest nonnegative real number M such that $\|Tx\| \le M\|x\|$ for each x in X.*

PROOF. Define $f\colon X \to \mathbb{R}$ by the formula $f(x) = \|Tx\|$. Since f is continuous, its supremum on the open unit ball of X is the same as its supremum on the closure B_X of that ball, which gives (a). If x is a nonzero element of B_X, then $\|x\|^{-1}x \in S_X$ and $f\big(\|x\|^{-1}x\big) = \|x\|^{-1}f(x) \ge f(x)$. It follows that the supremum of f on S_X is the same as its supremum on B_X whenever $X \neq \{0\}$, which gives (b). For (c), first notice that if $0 \le M < \|T\|$, then there is an x in B_X such that $\|Tx\| > M \ge M\|x\|$, so there is no nonnegative real number M smaller than $\|T\|$ such that $\|Tx\| \le M\|x\|$ for each x in X. If $x \in X$ and $x \neq 0$, then $\|x\|^{-1}\|Tx\| = \big\|T\big(\|x\|^{-1}x\big)\big\| \le \|T\|$, and so $\|Tx\| \le \|T\|\,\|x\|$. This last inequality is trivially true if $x = 0$, which finishes the proof of (c). ∎

So far, it has not been shown that the operator norm really is a norm. It is time to remedy this oversight.

1.4.8 Theorem. *Let X and Y be normed spaces. Then $B(X,Y)$ is a normed space under the operator norm. If Y is a Banach space, then so is $B(X,Y)$.*

PROOF. Suppose that $S, T \in B(X,Y)$. It is clear that $\|T\| \ge 0$. If $T \neq 0$, then there is an x_0 in X, necessarily nonzero, such that $Tx_0 \neq 0$, and so $T\big(\|x_0\|^{-1}x_0\big) \neq 0$. It follows that $T = 0$ if and only if $Tx = 0$ for each x in B_X, that is, if and only if $\|T\| = 0$. If α is a scalar, then

$$\|\alpha T\| = \sup\{\, \|\alpha Tx\| : x \in B_X \,\} = |\alpha| \sup\{\, \|Tx\| : x \in B_X \,\} = |\alpha|\,\|T\|.$$

If $x_0 \in B_X$, then

$$\|(S+T)(x_0)\| \le \|S\|\,\|x_0\| + \|T\|\,\|x_0\| \le \|S\| + \|T\|,$$

and so

$$\|S+T\| = \sup\{\, \|(S+T)(x)\| : x \in B_X \,\} \le \|S\| + \|T\|.$$

Thus, the operator norm is a norm on $B(X,Y)$.

Suppose that Y is a Banach space. Let (T_n) be a Cauchy sequence in $B(X,Y)$. If $x \in X$, then

$$\|T_nx - T_mx\| = \|(T_n - T_m)(x)\| \le \|T_n - T_m\|\,\|x\|$$

whenever $n, m \in \mathbb{N}$, from which it follows that the sequence $(T_n x)$ is Cauchy in Y and hence convergent. Define $T\colon X \to Y$ by the formula $Tx = \lim_n T_n x$. Because the vector space operations of Y are continuous, the map T is linear. To see that T is bounded, first notice that the boundedness of the Cauchy sequence (T_n) gives a nonnegative M such that $\|T_n\| \le M$ for each n, so that $\|T_n x\| \le M$ for each x in B_X and each n. Letting n tend to ∞ shows that $\|Tx\| \le M$ for each x in B_X, so T is bounded. To see that $\|T_n - T\| \to 0$ as $n \to \infty$, let ϵ be a positive number and let N_ϵ be a positive integer such that $\|T_n - T_m\| \le \epsilon$ whenever $n, m \ge N_\epsilon$. If $x \in B_X$ and $n, m \ge N_\epsilon$, then $\|T_n x - T_m x\| \le \|T_n - T_m\| \le \epsilon$. Keeping n fixed and letting m tend to ∞ shows that $\|T_n x - Tx\| \le \epsilon$ whenever $x \in B_X$ and $n \ge N_\epsilon$. Taking the supremum over all x in B_X then shows that $\|T_n - T\| \le \epsilon$ whenever $n \ge N_\epsilon$. It follows that $\lim_n T_n = T$, so $B(X, Y)$ is complete. ■

Henceforth, whenever X and Y are normed spaces and $B(X, Y)$ is also treated as a normed space, the operator norm is assumed unless specifically stated otherwise.

If X is a Banach space and Y is an incomplete normed space, then $B(X, Y)$ might not be a Banach space; see Exercise 1.41. In fact, it will follow from Exercise 1.117 that as long as X and Y are normed spaces and $X \ne \{0\}$, the space $B(X, Y)$ is a Banach space if *and only if* Y is a Banach space. If $X = \{0\}$, then $B(X, Y)$ contains only the zero operator and is trivially a Banach space, even if Y is not.

1.4.9 Proposition. *Let X and Y be normed spaces. If $T \in B(X, Y)$ and (T_n) is a sequence in $B(X, Y)$ such that $T_n \to T$, then $T_n x \to Tx$ for each x in X.*

PROOF. For each x in X,

$$\|T_n x - Tx\| = \|(T_n - T)(x)\| \le \|T_n - T\|\,\|x\| \to 0$$

as $n \to \infty$. ■

The converse of Proposition 1.4.9 is in general false, as should be expected. Convergence in $B(X, Y)$ represents *uniform* convergence on the closed unit ball of X, so it is not surprising that such convergence is not always implied by *pointwise* convergence on X. See Exercise 1.44.

Suppose that X, Y, and Z are normed spaces and that $S \in B(X, Y)$ and $T \in B(Y, Z)$. Then the product TS is in $B(X, Z)$ since the composite of continuous functions is continuous. It is too much to hope that $\|TS\|$ would have to equal $\|T\|\,\|S\|$; for example, it is easy to find nonzero members S and T of $B(\mathbb{R}^2)$ whose product is the zero operator. The following, at least, can be said.

1.4.10 Proposition. *Let X, Y, and Z be normed spaces. If $S \in B(X,Y)$ and $T \in B(Y,Z)$, then $TS \in B(X,Z)$ and $\|TS\| \le \|T\|\,\|S\|$.*

PROOF. As has already been mentioned above, the operator TS is continuous as the composite of two continuous functions. The continuity of TS also follows from the fact that for each x in X,

$$\|TS(x)\| = \|T(Sx)\| \le \|T\|\,\|Sx\| \le \|T\|\,\|S\|\,\|x\|.$$

This inequality also shows that $\|TS\| \le \|T\|\,\|S\|$, since $\|TS\|$ is the smallest nonnegative real number M such that $\|TS(x)\| \le M\|x\|$ for each x in X. ■

As was shown in Example 1.4.5, there are normed spaces X and Y such that some members of $L(X,Y)$ are unbounded. Of course, this cannot happen when $Y = \{0\}$, for then $L(X,Y)$ contains only the zero operator. This observation, together with the following two theorems, completely settles the question of which ordered pairs (X,Y) of normed spaces possess the property that $L(X,Y)$ has an unbounded member: this happens if and only if X is infinite-dimensional and $Y \ne \{0\}$.

1.4.11 Theorem. *Let X and Y be normed spaces such that X is infinite-dimensional and $Y \ne \{0\}$. Then some linear operator from X into Y is unbounded. In particular, every infinite-dimensional normed space has a linear functional on it that is unbounded.*

PROOF. Let (b_n) be a linearly independent sequence in S_X, let $\mathfrak{B}$ be a vector space basis for X that includes $\{\, b_n : n \in \mathbb{N} \,\}$, and let y be a nonzero member of Y. Define $T_{\mathfrak{B}} \colon \mathfrak{B} \to Y$ by letting $T_{\mathfrak{B}}(b_n) = ny$ for each n and letting $T_{\mathfrak{B}}(b) = 0$ for every other member b of $\mathfrak{B}$. Then $T_{\mathfrak{B}}$ can be extended to a member T of $L(X,Y)$, and T is not bounded since $T(S_X)$ is not a bounded set. ■

1.4.12 Theorem. *Let X and Y be normed spaces such that X is finite-dimensional. Then every linear operator from X into Y is bounded.*

PROOF. In this proof, the standard notation $\|\cdot\|$ is used for all norms except the special one about to be defined. Let V be the vector space underlying the normed space X and let n be the dimension of V. Since the theorem is trivial if $n = 0$, it can be assumed that $n \ge 1$. Let $x_1, \dots, x_n$ be a vector space basis for V. Define a norm $/\cdot/$ on V by the formula

$$/\alpha_1 x_1 + \cdots + \alpha_n x_n/ = |\alpha_1| + \cdots + |\alpha_n|,$$

and let W be the normed space $(V, /\cdot/)$. It follows immediately from the definition of $/\cdot/$ that a sequence $(\alpha_1^{(j)} x_1 + \cdots + \alpha_n^{(j)} x_n)$ in W converges to some member $\alpha_1 x_1 + \cdots + \alpha_n x_n$ of W if and only if $\lim_j \alpha_m^{(j)} = \alpha_m$ when $m = 1, \dots, n$.

Claim: If Z is a normed space, then each member of $L(W, Z)$ is bounded. To see this, suppose that $T \in L(W, Z)$ and $\alpha_1 x_1 + \cdots + \alpha_n x_n \in W$. Then

$$\begin{aligned}\|T(\alpha_1 x_1 + \cdots + \alpha_n x_n)\| &\leq |\alpha_1|\,\|Tx_1\| + \cdots + |\alpha_n|\,\|Tx_n\| \\ &\leq \big(\|Tx_1\| + \cdots + \|Tx_n\|\big) \cdot /\alpha_1 x_1 + \cdots + \alpha_n x_n/,\end{aligned}$$

so T is bounded. This proves the claim.

Let I be the identity operator on V, viewed as a member of $L(X, W)$. Since every T in $L(X, Y)$ can be written in the form $T_W I$ where T_W is just T viewed as a member of $L(W, Y)$, the theorem will be proved once it is shown that I is bounded. To this end, let $(\alpha_1^{(j)} x_1 + \cdots + \alpha_n^{(j)} x_n)$ be a sequence in S_W. Then $(\alpha_1^{(j)}), \ldots, (\alpha_n^{(j)})$ are bounded sequences of scalars, so there are scalars $\alpha_1, \ldots, \alpha_n$ and an increasing sequence (j_k) of positive integers such that $\lim_k \alpha_m^{(j_k)} = \alpha_m$ when $m = 1, \ldots, n$. It follows that

$$\lim_k (\alpha_1^{(j_k)} x_1 + \cdots + \alpha_n^{(j_k)} x_n) = \alpha_1 x_1 + \cdots + \alpha_n x_n$$

and that $\alpha_1 x_1 + \cdots + \alpha_n x_n \in S_W$. Therefore every sequence in S_W has a subsequence converging to a member of S_W, so S_W is compact. Since $I^{-1} \in L(W, X)$, the function $w \mapsto \|I^{-1} w\|$ from W into $\mathbb{R}$ is continuous and so attains a *positive* minimum on the compact set S_W. If I were not bounded, then there would be a sequence (z_j) in B_X such that $/Iz_j/ \geq j$ for each j. Letting $w_j = /Iz_j/^{-1} Iz_j$ for each j yields a sequence (w_j) in S_W such that $\|I^{-1} w_j\| = /Iz_j/^{-1} \|z_j\| \to 0$ as $j \to \infty$, a contradiction. Thus, the operator I must be bounded. ■

An important technique was used in the proof of the preceding theorem. To prove that a linear operator T from X into Y must be continuous, it was shown that there is a normed space W such that all the linear operators from W into Y have this property, and that T can be factored as the product of a continuous linear operator from X into W and a linear operator from W into Y. Variations of this technique with other properties besides continuity have many applications in Banach space theory.

1.4.13 Definitions. Suppose that T is a linear operator from a normed space X into a normed space Y. Then T is an *isomorphism* or *normed space isomorphism* into Y if it is one-to-one and continuous and its inverse mapping T^{-1} is continuous on the range of T. The operator T is an *isometric isomorphism* or *linear isometry* if $\|Tx\| = \|x\|$ whenever $x \in X$. The space X is *embedded* in Y if there is an isomorphism from X into Y, and is *isometrically embedded* in Y if there is an isometric isomorphism from X into Y. The spaces X and Y are *isomorphic* if there is an isomorphism from X onto Y, and are *isometrically isomorphic* if there is an isometric isomorphism from X onto Y. If X and Y are isomorphic, then this is denoted by writing $X \cong Y$.

Several points about the preceding definitions are worth emphasizing. In this book an isomorphism T is a one-to-one linear mapping from a normed space X *into* a normed space Y that is also a homeomorphism from X onto the range of T. This differs from the usual convention that a mapping called an isomorphism from an object A into an object B should be onto B, but agrees with normed space custom. A normed space isomorphism $T\colon X \to Y$ is essentially a mapping that provides a way of identifying both the vector space structure and the topology of X with those of $T(X)$. An isometric isomorphism does this while also identifying the norms of X and $T(X)$.

1.4.14 Proposition. *Let T be a linear operator from a normed space X into a normed space Y.*

(a) *The operator T is an isomorphism if and only if there are positive constants s and t such that $s\|x\| \le \|Tx\| \le t\|x\|$ whenever $x \in X$.*

(b) *If T is an isometric isomorphism, then T is an isomorphism.*

(c) *If X is a Banach space and T is an isomorphism, then $T(X)$ is a Banach space.*

PROOF. The proposition is trivially true if $X = \{0\}$, so it can be assumed that $X \neq \{0\}$. It follows that T and T^{-1} are both nonzero whenever T is an isomorphism, giving them both positive norms. For (a), notice that if T is an isomorphism, then $\|Tx\| \le \|T\|\,\|x\|$ and $\|x\| = \|T^{-1}(Tx)\| \le \|T^{-1}\|\,\|Tx\|$ for each x in X, so letting $s = \|T^{-1}\|^{-1}$ and $t = \|T\|$ gives the desired inequalities. Suppose conversely that there are positive real numbers s and t such that $s\|x\| \le \|Tx\| \le t\|x\|$ for each x in X. Then the second half of the inequality shows that T is bounded while the first half shows that $Tx \neq 0$ whenever $x \neq 0$, that is, that T is one-to-one. Since $\|T^{-1}(Tx)\| = \|x\| \le s^{-1}\|Tx\|$ for each x in X, the operator T^{-1} is bounded, so T is an isomorphism.

Part (b) follows immediately from (a). For (c), suppose that X is a Banach space and T is an isomorphism. Let (y_n) be a Cauchy sequence in $T(X)$. Since there is a positive number s such that $s\|T^{-1}y\| \le \|y\|$ for each y in $T(X)$, it follows that $(T^{-1}y_n)$ is a Cauchy sequence in X and so converges to some x_0. The continuity of T implies that (y_n) converges to Tx_0, proving that $T(X)$ is complete. ∎

The significance of part (c) of the preceding proposition comes from the fact that, in general, homeomorphisms between metric spaces do not have to preserve completeness. See Exercise 1.42.

An easy application of Theorem 1.4.12 immediately produces a large supply of isomorphic normed spaces.

1.4.15 Theorem. *Let n be a nonnegative integer and let X and Y be n-dimensional normed spaces over $\mathbb{F}$. Then every linear operator from X onto Y is an isomorphism.*

PROOF. Let T be a linear operator from X onto Y. Since X and Y have the same finite dimension, the operator T is one-to-one. By Theorem 1.4.12, both T and T^{-1} are bounded. ■

1.4.16 Corollary. *Let n be a nonnegative integer. Then all n-dimensional normed spaces over $\mathbb{F}$ are isomorphic to each other.*

PROOF. Let X and Y be n-dimensional normed spaces over $\mathbb{F}$. It is a standard fact from linear algebra that there is a linear operator from X onto Y; recall that such an operator can be found by letting $\mathfrak{B}_X$ and $\mathfrak{B}_Y$ be vector space bases for X and Y respectively, letting f be any map from $\mathfrak{B}_X$ onto $\mathfrak{B}_Y$, and then letting T_f be a linear extension of f to all of X. By the theorem, each such linear operator is a normed space isomorphism. ■

1.4.17 Corollary. *Every finite-dimensional vector space has exactly one norm topology.*

PROOF. Let X be a vector space of finite dimension n and let T be a linear operator from X onto Euclidean n-space. It is easily checked that the formula $\|x\|_T = \|Tx\|$ defines a norm on X. If $\|\cdot\|_0$ is any other norm on X, then the identity operator on X, viewed as a linear operator from $(X, \|\cdot\|_T)$ onto $(X, \|\cdot\|_0)$, is an isomorphism, so the two norms induce the same topology. ■

1.4.18 Corollary. *For each nonnegative integer n, the only norm topology that $\mathbb{F}^n$ can have is its Euclidean topology.*

Of course, the preceding two corollaries do not prevent a finite-dimensional vector space from having many different norms, as Example 1.2.9 shows. However, the different norms must all induce the same topology.

The first of the following two corollaries is obtained by observing that each normed space with finite dimension n must be isomorphic to Euclidean n-space, a Banach space. The second then follows from the first, since each complete subset of a metric space is closed in that space.

1.4.19 Corollary. *Every finite-dimensional normed space is a Banach space.*

1.4.20 Corollary. *Every finite-dimensional subspace of a normed space is a closed subset of the space.*

1.4.21 Corollary. *Every finite-dimensional normed space has the* Heine-Borel *property, that is, the property that all closed bounded subsets of the space are compact.*

PROOF. Let n be a nonnegative integer. Notice first that if $n > 0$, then Euclidean n-space $\mathbb{F}^n$ has the Heine-Borel property, for if A is a closed bounded subset of $\mathbb{F}^n$ and $\left((\alpha_1^{(j)},\dots,\alpha_n^{(j)})\right)$ is a sequence in A, then the boundedness of each of the sequences $(\alpha_m^{(j)})$ such that $m = 1,\dots,n$ assures the existence of a subsequence (j_k) of $\mathbb{N}$ and scalars $\alpha_1,\dots,\alpha_n$ such that $(\alpha_m^{(j_k)}) \to \alpha_m$ for each m, which implies that $(\alpha_1^{(j_k)},\dots,\alpha_n^{(j_k)}) \to (\alpha_1,\dots,\alpha_n)$. It is trivially true that $\mathbb{F}^0$ has the Heine-Borel property.

Let X be a normed space having finite dimension n and let T be an isomorphism from X onto Euclidean n-space. If S is a closed bounded subset of X, then $T(S)$ is closed and bounded in $\mathbb{F}^n$ and therefore compact, so S is itself compact. ■

A little more work shows that the Heine-Borel property actually characterizes the finite-dimensional normed spaces among all normed spaces.

1.4.22 Lemma. *If X is an infinite-dimensional normed space, then there is a linearly independent sequence (x_n) in S_X such that $\|x_n - x_m\| \geq 1$ whenever $n \neq m$.*

PROOF. The sequence (x_n) is constructed inductively, beginning with an arbitrary element x_1 of S_X. Suppose that $n \geq 2$ and that linearly independent elements $x_1,\dots,x_{n-1}$ of S_X have been found such that $\|x_j - x_k\| \geq 1$ if $j,k \leq n-1$ and $j \neq k$. Let $Y = \langle x_1,\dots,x_{n-1}\rangle$, a finite-dimensional subspace of X. The induction will be complete once a member x_n of S_X is found such that $d(x_n, Y) = 1$, for then $x_1,\dots,x_n$ must be linearly independent and $\|x_n - x_m\| \geq 1$ if $m < n$. Fix an element z of $X \setminus Y$ and let $t = d(z, Y)$, a positive number since Y is closed. A subspace of a vector space remains unchanged when multiplied by a nonzero scalar or translated by one of its own elements, which in particular implies that $t^{-1}Y = Y$. It follows that $1 = t^{-1}d(z,Y) = d(t^{-1}z, t^{-1}Y) = d(t^{-1}z, Y)$. Let (y_j) be a sequence in Y such that $\|y_j - t^{-1}z\| \to 1$. Since (y_j) is a bounded sequence in Y and Y has the Heine-Borel property, the sequence (y_j) has a subsequence (y_{j_k}) that converges to some y in Y. Then $\|y - t^{-1}z\| = \lim_k \|y_{j_k} - t^{-1}z\| = 1$ and $d(y - t^{-1}z, Y) = d(y - t^{-1}z, y - Y) = d(t^{-1}z, Y) = 1$. Let $x_n = y - t^{-1}z$ to complete the induction and the proof. ■

1.4.23 Theorem. (F. Riesz, 1918 [195]). *A normed space X is finite-dimensional if and only if it has the Heine-Borel property, which happens if and only if S_X is compact.*

PROOF. If X is finite-dimensional, then it has the Heine-Borel property, so the closed bounded set S_X is compact. If X is infinite-dimensional, then the lemma produces a sequence in S_X with no convergent subsequence, so S_X is not compact and X lacks the Heine-Borel property. ■

Exercises

1.40 Supply the details in Example 1.4.6.

1.41 Suppose that Y is the vector space of finitely nonzero sequences, equipped with the ℓ_1 norm. Show that $B(\ell_1, Y)$ is not a Banach space.

1.42 Define $d\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ by the formula $d(x,y) = |\tan^{-1}(x)-\tan^{-1}(y)|$, where $\tan^{-1}$ denotes the usual single-valued inverse tangent function from $\mathbb{R}$ onto $(-\pi/2, \pi/2)$. Prove that d is an incomplete metric on $\mathbb{R}$ that induces the usual topology of $\mathbb{R}$. Conclude that it is possible for an incomplete metric space to be homeomorphic to a complete one.

1.43 Suppose that T is a linear operator from a normed space X into a normed space Y such that $\sum_n Tx_n$ is a convergent series in Y whenever $\sum_n x_n$ is an absolutely convergent series in X. Prove that T is bounded.

1.44 For each positive integer j, define $e_j^* : c_0 \to \mathbb{F}$ by the formula $e_j^*(\alpha_n) = \alpha_j$. Show that (e_j^*) is a sequence in $B(c_0, \mathbb{F})$ that does not converge to the zero operator $\mathbf{0}$ of $B(c_0, \mathbb{F})$, even though $\lim_j e_j^*(\alpha_n) = \mathbf{0}(\alpha_n)$ whenever $(\alpha_n) \in c_0$.

1.45 Prove that no Banach space has a countably infinite vector space basis. (Suppose that X is a normed space with a countably infinite vector space basis (b_n). For each positive integer n, let $Y_n = \langle b_1, \ldots, b_n\rangle$. Select an x_1 in Y_1 such that $\|x_1\| = 1$. Use an argument like that of Lemma 1.4.22 to select an x_2 in Y_2 such that $d(x_2, Y_1) = \|x_2 - x_1\| = 1/4$. Select an x_3 in Y_3 such that $d(x_3, Y_2) = \|x_3 - x_2\| = 1/16$. Continuing in this vein yields a Cauchy sequence that cannot converge to anything in any Y_n.)

1.46 Suppose that Y is a finite-dimensional normed space. Let T be a linear operator from a normed space X onto Y, where X is not assumed to be finite-dimensional and T is not assumed to be bounded. Prove that T is an *open mapping*, that is, that $T(U)$ is an open subset of Y whenever U is an open subset of X. (Notice that it is enough to prove that $T(B_X)$ includes a neighborhood of 0.)

1.47 Let f be an unbounded linear functional on a normed space X. Prove that if U is a nonempty open subset of X, then $f(U) = \mathbb{F}$.

1.48 Let X and Y be normed spaces and let $f\colon X \to Y$ be *additive*; that is, let $f(x_1 + x_2) = f(x_1) + f(x_2)$ whenever $x_1, x_2 \in X$.

(a) Prove that $f(rx) = rf(x)$ for each rational number r and each x in X.

(b) Prove that if X and Y are real normed spaces and f is continuous, then f is linear.

(c) Show that the conclusion of (b) might not hold if X and Y are complex normed spaces and f is continuous.

1.49 Let c be the Banach space of convergent sequences of scalars defined in Exercise 1.25.

(a) Prove that c is isomorphic to c_0.

(b) Suppose that $x \in S_{c_0}$. Prove that there exist x_1 and x_2 in S_{c_0} such that $x_1 \neq x_2$ and $x = \frac{1}{2}(x_1 + x_2)$.

(c) Conclude from (b) that c is not isometrically isomorphic to c_0.

1.50 The purpose of this exercise is to show that, unlike finite-dimensional vector spaces, infinite-dimensional vector spaces never have unique norm topologies. Let $(X, \|\cdot\|)$ be an infinite-dimensional normed space.

(a) Construct an unbounded one-to-one linear operator T from $(X, \|\cdot\|)$ onto itself.

(b) Let $\|x\|_T = \|Tx\|$ whenever $x \in X$. Show that $\|\cdot\|_T$ is a norm, that T is an isometric isomorphism from $(X, \|\cdot\|_T)$ onto $(X, \|\cdot\|)$, and that $\|\cdot\|_T$ is a Banach norm if and only if $\|\cdot\|$ is a Banach norm.

(c) Show that the topologies induced by $\|\cdot\|$ and $\|\cdot\|_T$ are different.

1.51 Here is another proof of Corollary 1.4.18 that has a more geometric flavor. Let n be a nonnegative integer. Let $\|\cdot\|_0$ be a norm on $\mathbb{F}^n$, let $\mathfrak{T}$ be the topology induced by this norm, and let U_0 and U_e be the open unit balls for $\|\cdot\|_0$ and the Euclidean norm respectively. In the following, any term preceded by "e-" refers to the Euclidean topology or norm. Do not use Theorem 1.4.12 or any results based on it in your arguments. You may use the fact that Euclidean n-space has the Heine-Borel property, since the demonstration of that fact in the proof of Corollary 1.4.21 is elementary.

(a) Using the fact that e-convergence implies coordinatewise convergence when $n \geq 1$, show that each e-convergent sequence in $\mathbb{F}^n$ is $\mathfrak{T}$-convergent to the e-limit. Conclude that every $\mathfrak{T}$-open subset of $\mathbb{F}^n$ is e-open.

(b) Let C be an e-closed e-unbounded convex set in $\mathbb{F}^n$ that contains the origin. Show that C includes a ray emanating from the origin, that is, a set of the form $\{\, tx : t \geq 0 \,\}$ where x is a nonzero element of $\mathbb{F}^n$.

(c) Let C be an e-open e-unbounded convex set in $\mathbb{F}^n$ that contains the origin. Show that C includes a ray emanating from the origin.

(d) Show that $sU_0 \subseteq U_e$ for some positive s. Conclude that every e-open subset of $\mathbb{F}^n$ is $\mathfrak{T}$-open, and that $\mathfrak{T}$ must therefore be the Euclidean topology. This concludes the proof of Corollary 1.4.18.

(e) Obtain Corollary 1.4.16 from Corollary 1.4.18. (Let T be a linear operator from an n-dimensional normed space X onto $\mathbb{F}^n$ and define a norm $\|\cdot\|_0$ on $\mathbb{F}^n$ by letting $\|y\|_0 = \|T^{-1}y\|$. Use this norm to show that X is isomorphic to Euclidean n-space.)

1.5 Baire Category

This section could have been marked optional, since the material developed in it is used only in optional Section 2.3. Though it is useful to have some form of the main result of this section, the Baire category theorem, to derive the three fundamental theorems to be presented in the next section, a weak form of the Baire category theorem already obtained as Theorem 1.3.14 is enough to do that. However, the reader who is not already familiar with

Baire category should spend some time with this section, since Baire category is an important component of the working analyst's toolkit.

It is often useful to be able to prove that a set is too large to be written as a union of countably many other sets that are in some sense very small. For example, facts about the real line that depend on its uncountability are true because the real line is "large" enough that it cannot be written as a union of countably many one-point "very small" sets. Similarly, many results in measure theory are obtained from knowing that a set of nonzero measure is too "large" to be a union of countably many "very small" sets of measure zero.

The concept of Baire category, due to Louis René Baire, gives a topological meaning to the notion of the size of a set. Baire's approach is based on density. A subset A of a topological space X is considered to be very small in Baire's sense if there is no nonempty open subset U of X such that $A \cap U$ is dense in U, that is, if $\overline{A}$ has empty interior. Baire's large sets are those that are not unions of countably many of these very small sets.

1.5.1 Definitions. Let X be a topological space. A subset of X is

(a) *nowhere dense in* X if its closure has empty interior;

(b) *of the first category in* X or *meager in* X if it is the union of a countable collection of sets that are nowhere dense in X;

(c) *of the second category in* X or *nonmeager in* X if it is not of the first category in X.

In particular, the set A is of the *first* or *second category in itself* if A has that category as a subset of the topological space $(A, \mathfrak{T}_A)$, where $\mathfrak{T}_A$ is the relative topology that A inherits from X.

1.5.2 Example. If α is any real number, then the set $\{\alpha\}$ is obviously of the first category, in fact nowhere dense, in the real line, and is equally obviously of the second category in itself. Since the set $\mathbb{Q}$ of rational numbers can be written as the union of countably many singleton subsets of $\mathbb{R}$, it follows that $\mathbb{Q}$ is of the first category in $\mathbb{R}$ while being at the same time dense in $\mathbb{R}$.

Here are some simple but useful facts about Baire category, followed by the main result of this section.

1.5.3 Proposition. *Let X be a topological space.*

(a) *If a subset A of X is nowhere dense or of the first category in X, then every subset of A has that same property.*

(b) *A subset of X is nowhere dense in X if and only if the interior of its complement is dense in X.*

(c) *A subset of X is of the first category in X if and only if its complement is the intersection of a countable collection of sets each of whose interiors is dense in X.*

(d) *The union of a countable collection of sets each of the first category in X is of the first category in X.*

(e) *If a subset of X is of the second category in X, then it is of the second category in itself.*

PROOF. Clearly, subsets of nowhere dense subsets of X are themselves nowhere dense in X. It follows that if A is the union of a countable collection $\mathfrak{C}$ of nowhere dense subsets of X and $B \subseteq A$, then B is the union of the countable collection $\{ B \cap C : C \in \mathfrak{C} \}$ of nowhere dense subsets of X, which proves (a).

Now suppose that A is an arbitrary subset of X. Then $(\overline{A})^\circ = \varnothing$ if and only if $X \setminus \overline{A}$ intersects every nonempty open subset of X, which happens if and only if $X \setminus \overline{A}$ is dense in X. Since $X \setminus \overline{A} = (X \setminus A)^\circ$, this proves (b). Part (c) then follows from a straightforward application of De Morgan's laws.

Part (d) is obvious. For (e), notice that if $A \subseteq X$, then a subset of A that is nowhere dense in A is nowhere dense in X, from which it follows that if A is of the first category in itself, then it is also of the first category in X. ■

1.5.4 The Baire Category Theorem. (L. R. Baire, 1899 [9]). *Let X be a nonempty complete metric space.*

(a) *If $\mathfrak{U}$ is a countable collection of open subsets of X each of which is dense in X, then $\bigcap\{ U : U \in \mathfrak{U} \}$ is dense in X.*

(b) *Every nonempty open subset of X is of the second category in X and hence in itself. In particular, the entire space X is of the second category in itself.*

PROOF. For (a), suppose that $\mathfrak{U}$ is a countable collection of dense open subsets of X. It can be assumed that $\mathfrak{U} \neq \varnothing$. Then the members of $\mathfrak{U}$ can be written as a sequence (U_n), which might require that some member of $\mathfrak{U}$ be repeated infinitely often in the list. Let U be a nonempty open subset of X. To prove (a), it is enough to show that $U \cap (\bigcap_n U_n) \neq \varnothing$. Since U_1 is dense in X, there is a closed ball B_1 of radius no more than 1 included in $U \cap U_1$. Now suppose that $m \geq 2$ and that closed balls $B_1, \dots, B_{m-1}$ have been found such that these conditions are satisfied when $1 \leq j \leq m-1$:

(1) $B_{j-1} \supseteq B_j$ if $j \geq 2$;

(2) $B_j \subseteq U \cap U_1 \cap \cdots \cap U_j$;

(3) the radius of B_j is no more than j^{-1}.

Let G_{m-1} be the open ball with the same center and radius as B_{m-1}. Then the density of U_m in X implies that there is a closed ball B_m of radius no more than m^{-1} included in $G_{m-1} \cap U_m$. It follows by induction that there is a sequence (B_n) of closed balls such that (1), (2), and (3) are satisfied for each positive integer j. The centers of these closed balls form a Cauchy sequence that converges to some x in X. This x must lie in each B_n and therefore in U and each U_n, which shows that $U \cap (\bigcap_n U_n) \neq \emptyset$ and proves (a).

For (b), notice that the complement of a nonempty open subset of X obviously cannot be dense in X, and therefore, by (a), cannot be the intersection of a countable collection of sets each of whose interiors is dense in X. It follows from Proposition 1.5.3 (c) that every nonempty open subset of X is of the second category in X, and therefore of the second category in itself by Proposition 1.5.3 (e). ■

Of course, the metric space formed from the empty set with its only possible metric is a complete metric space of the first category in itself. Nonempty incomplete metric spaces can be of either the first or second category in themselves. The rationals with the metric inherited from the reals form an incomplete metric space of the first category in itself, while the interval $(0, 1)$ in $\mathbb{R}$ with the metric inherited from $\mathbb{R}$ is an incomplete metric space that is of the second category in itself by part (b) of the Baire category theorem.

Suppose that P is a property defined for the elements of some nonempty complete metric space X. It follows from the Baire category theorem that one way to prove that some member of X has property P is to show that the members of X lacking property P form a set of the first category in X. If this can be done, then the collection of elements of X having property P is a set of the second category in X whose complement is of the first category in X, so it can even be said that in some sense the "typical" member of X has property P. The following is a classical application of this idea.

From Newton's time through the early part of the nineteenth century, most mathematicians assumed that a continuous real-valued function defined on an interval in the real line must be differentiable over most of its domain. In 1834, Bernhard Bolzano gave an example of a real-valued function continuous on an interval though differentiable nowhere on that interval,[4] but for almost a century afterward mathematicians treated such functions as pathological. However, in 1931 Stefan Banach showed that, in a sense, the vast majority of continuous scalar-valued functions whose domain is a given interval in $\mathbb{R}$ are not differentiable anywhere.

1.5.5 Theorem. (S. Banach, 1931 [12]). *Let D_+ be the collection of all members f of $C[0, 1]$ for which there is a point x_f in $[0, 1)$ at which f has a finite right-hand derivative. Then D_+ is of the first category in $C[0, 1]$.*

---

[4]Weierstrass is usually given credit for finding the first everywhere continuous, nowhere differentiable function, but his example was first presented in lectures in 1861 and in a paper to the Berlin Academy in 1872. See pages 577 and 627 of [38] for a discussion of Bolzano's priority.

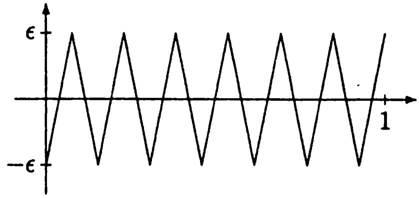


FIGURE 1.1. A member of $C[0,1]$ with norm ϵ and large right-hand derivative.

PROOF. For each positive integer n, let U_n be the collection of all members f of $C[0,1]$ such that for each x in $[0,1-n^{-1}]$,

$$\sup\left\{\left|\frac{f(y)-f(x)}{y-x}\right| : x < y < x + \frac{1}{n}\right\} > n.$$

It will be shown that each U_n is a dense open subset of $C[0,1]$. To this end, fix an n_0 in $\mathbb{N}$. Suppose that (f_m) is a sequence in $C[0,1] \setminus U_{n_0}$ converging to some f_0 in $C[0,1]$. To show that U_{n_0} is open, it is enough to show that $f_0 \in C[0,1] \setminus U_{n_0}$. For each positive integer m, let x_m be an element of $[0, 1-n_0^{-1}]$ such that $|(f_m(y) - f_m(x_m))/(y - x_m)| \leq n_0$ whenever $x_m < y < x_m + n_0^{-1}$. By thinning the sequence (f_m) if necessary, it can be assumed that there is an x_0 in $[0, 1-n_0^{-1}]$ such that $x_m \to x_0$. If y is such that $x_0 < y < x_0 + n_0^{-1}$, then for m large enough that $x_m < y < x_m + n_0^{-1}$,

$$n_0 \geq \left|\frac{f_m(y) - f_m(x_m)}{y - x_m}\right| \to \left|\frac{f_0(y) - f_0(x_0)}{y - x_0}\right| \quad \text{as } m \to \infty,$$

and so $|(f_0(y) - f_0(x_0))/(y - x_0)| \leq n_0$. It follows that $f_0 \in C[0,1] \setminus U_{n_0}$, and therefore that U_{n_0} is open.

The proof that U_{n_0} is dense in $C[0,1]$ uses the Weierstrass approximation theorem, which says that the polynomials are dense in $C[0,1]$; see, for example, [19] or [201]. If ϵ and M are positive numbers, then Figure 1.1 shows how to construct a sawtooth function that is in $C[0,1]$, has norm ϵ, and has a right-hand derivative with absolute value greater than M at each point of $[0,1)$. Since each polynomial on $[0,1]$ has a bounded right-hand derivative on $[0,1)$, it follows that if p is a polynomial on $[0,1]$ and $\epsilon > 0$, then a sawtooth function can be added to p to yield a member u of U_{n_0} such that $\|u - p\|_\infty \leq \epsilon$. The set U_{n_0} is therefore dense in $C[0,1]$.

Thus, each U_n is a dense open subset of $C[0,1]$, and so, by Proposition 1.5.3 (c), the set $C[0,1] \setminus (\bigcap_n U_n)$ is of the first category in $C[0,1]$. Since a member of $C[0,1]$ with a finite right-hand derivative at some point of $[0,1)$ cannot lie in every U_n, the set of all members of $C[0,1]$ that are differentiable from the right anywhere on $[0,1)$ is included in $C[0,1] \setminus (\bigcap_n U_n)$ and therefore is of the first category in $C[0,1]$. ∎

1.5.6 Corollary. *Let D be the collection of all members f of $C[0,1]$ for which there is a point x_f in $(0,1)$ at which f has a finite derivative. Then D is of the first category in $C[0,1]$.*

Since the Baire category theorem prevents the nonempty complete metric space $C[0,1]$ from being the union of two sets of the first category in $C[0,1]$, the next corollary follows immediately from the preceding one.

1.5.7 Corollary. *The collection of all members of $C[0,1]$ that are not differentiable anywhere on $(0,1)$ is of the second category in $C[0,1]$.*

Exercise 1.45 at the end of the preceding section asks for a proof that Banach spaces never have countably infinite vector space bases. The argument outlined in that exercise involves the careful construction of a certain nonconvergent Cauchy sequence. The result can also be obtained very easily from the Baire category theorem.

1.5.8 Theorem. *No Banach space has a countably infinite vector space basis.*

PROOF. Suppose that X is a normed space with a countably infinite vector space basis $\{x_n : n \in \mathbb{N}\}$. If U is an open subset of X and $x \in U$, then $X = \bigcup_{n=1}^{\infty} n(-x+U)$, and so $X = \langle U \rangle$. It follows that no finite-dimensional subspace of X includes a nonempty open subset of X.

For each positive integer n, let $F_n = \langle x_1, \dots, x_n \rangle$. Then each F_n is closed by Corollary 1.4.20 and includes no nonempty open subset of X, and so is nowhere dense in X. Since $X = \bigcup_n F_n$, the space X is of the first category in itself and therefore is not complete. ■

Readers interested in the analogies and interplay between measure and Baire category should become acquainted with J. C. Oxtoby's book [178]. R. P. Boas's monograph [31], especially its Section 10, is another good place to look for additional interesting applications of the Baire category theorem.

Exercises

1.52 Derive Theorem 1.3.14 from the Baire category theorem.

1.53 Prove that the union of finitely many nowhere dense subsets of a topological space must itself be nowhere dense in the space.

1.54 Prove that a topological space is of the first category in itself if and only if it is the union of countably many closed sets each having empty interior.

1.55 Prove that if X_1 and X_2 are topological spaces at least one of which is of the first category in itself, then the topological product $X_1 \times X_2$ is of the first category in itself.

1.56 The closure of a nowhere dense set is obviously nowhere dense. Must the closure of a set of the first category in a topological space X be of the first category in X?

1.57 Prove that the boundary of a closed subset of a topological space must be nowhere dense in the space. Conclude that a subset of a topological space is nowhere dense in that space if and only if it is a subset of the boundary of some closed set.

1.58 A subset A of a topological space X is *nearly open* if there are subsets M_1 and M_2 of the first category in X such that $(A \setminus M_1) \cup M_2$ is open. Prove that every Borel subset of a topological space is nearly open. (The preceding exercise might be helpful.)

1.59 An element x of a subset A of a metric space is an *isolated point* of A if there is a positive ϵ such that no point of A besides x itself has distance to x less than ϵ. A subset of a metric space is *perfect* if it is closed and has no isolated points. Prove that every nonempty perfect subset of a complete metric space is uncountable. Use this to give a proof that the reals are uncountable without using the usual diagonalization argument. Prove in a similar way that the Cantor set is uncountable.

1.60 Prove that the rational numbers are not a G_δ subset of the reals. (Recall that a G_δ set is a set that is the intersection of countably many open sets.)

1.61 Suppose that f is any function from $\mathbb{R}$ into $\mathbb{R}$. Prove that the points of continuity of f form a G_δ subset of $\mathbb{R}$. Conclude that no function from $\mathbb{R}$ into $\mathbb{R}$ is continuous precisely on the rationals. Is the same true for the irrationals? (The preceding exercise may help.)

1.62 Show that the real line can be partitioned into two subsets, one of which is of the first category in $\mathbb{R}$ and the other of which has Lebesgue measure zero. Thus, though the real line is large in both Baire's topological sense and Lebesgue's measure-theoretic sense, it can be written as the disjoint union of a topologically small set and a measure-theoretically small set.

1.63 Prove that every nonempty locally compact Hausdorff space is of the second category in itself.

1.6 Three Fundamental Theorems

Much of the theory of Banach spaces is based on three related results, called the open mapping theorem, uniform boundedness principle, and closed graph theorem, whose conclusions do not hold for arbitrary normed spaces. The purpose of this section is to obtain these fundamental theorems. The general plan of attack is to derive each of them from a result called Zabreĭko's lemma that is itself a straightforward consequence of Theorem 1.3.14. Other interesting proofs of these three theorems can be found in many analysis texts. See, for example, [67] and [200], as well as Section 2.3 of this book, all of which contain more general versions of these three results that are proved without using Zabreĭko's lemma.

1.6.1 Definition. A *seminorm* or *prenorm* on a vector space X is a real-valued function p on X such that the following conditions are satisfied by all members x and y of X and each scalar α:

(1) $p(\alpha x) = |\alpha|\, p(x)$;

(2) $p(x + y) \le p(x) + p(y)$.

For example, if X and Y are normed spaces and T is a linear operator from X into Y, then the function $x \mapsto \|Tx\|$ from X into $\mathbb{R}$ is easily seen to be a seminorm on X, called the *seminorm induced by* T. Of course, a norm is always a seminorm, and is in fact just the seminorm induced by the identity operator on the space.

Suppose that p is a seminorm on a vector space X. Then $p(0) = 0$, since $p(0) = p(0 \cdot 0) = 0 \cdot p(0)$; notice that some of those zeros are scalars and some are vectors. It then follows that p is nonnegative-real-valued rather than just real-valued, since $0 = p(0) \le p(x) + p(-x) = 2p(x)$ for each x in X. A seminorm is therefore a function on a vector space that satisfies the definition of a norm, except that the value of the seminorm of a nonzero vector is allowed to be zero.

1.6.2 Definition. A function f from a normed space X into the nonnegative reals is *countably subadditive* if $f(\sum_n x_n) \le \sum_n f(x_n)$ for each convergent series $\sum_n x_n$ in X.

For example, the norm of a normed space is always countably subadditive by Proposition 1.3.7 (e). More generally, let p be a seminorm on a normed space X. If p is continuous, then p is countably subadditive, as can be seen by letting $\sum_n x_n$ be a convergent series in X and letting m tend to infinity in the inequality $p(\sum_{n=1}^m x_n) \le \sum_{n=1}^m p(x_n)$. Conversely, if p is countably subadditive, then p must be continuous *provided that X is a Banach space.* That is the content of Zabreĭko's lemma.

1.6.3 Zabreĭko's Lemma. (P. P. Zabreĭko, 1969 [246]). *Every countably subadditive seminorm on a Banach space is continuous.*

PROOF. Let p be a countably subadditive seminorm on a Banach space X. If p is continuous at 0 and x is an element of X, then the argument used in the proof of Proposition 1.3.1, with the norm function replaced by p, shows that $|p(x) - p(y)| \le p(x - y) = |p(x - y) - p(0)|$ whenever $y \in X$, which implies the continuity of p at x. Thus, Zabreĭko's lemma will be proved once it is shown that p is continuous at 0.

Let $G = \{x : x \in X, p(x) < 1\}$, the "open unit ball" for p. If $t > 0$, then $tG = \{x : x \in X, p(x) < t\}$ by an application of property (1) in Definition 1.6.1, and so $x \in tG$ whenever $x \in X$ and $t > p(x)$. Thus, the set G is absorbing. If $x, y \in G$ and $0 < t < 1$, then

$$p\big(tx + (1-t)y\big) \le tp(x) + (1-t)p(y) < 1,$$

so $tx + (1-t)y \in G$, which shows that G is convex. Therefore $\overline{G}$ is closed, convex, and absorbing, and so by Theorem 1.3.14 includes an open ball U centered at 0 with some positive radius ϵ. If there is a positive real number s such that $p(x) < s$ whenever $\|x\| < \epsilon$, then $p(x) < t$ whenever $t > 0$ and $\|x\| < s^{-1}t\epsilon$, which would imply the continuity of p at 0. Thus, the proof will be complete once such an s is found.

Fix an x in X such that $\|x\| < \epsilon$. Since $x \in U \subseteq \overline{G}$, there is an x_1 in G such that $\|x - x_1\| < 2^{-1}\epsilon$. Since

$$x - x_1 \in 2^{-1}U \subseteq 2^{-1}\overline{G} = \overline{2^{-1}G},$$

there is an x_2 in $2^{-1}G$ such that $\|x - x_1 - x_2\| < 2^{-2}\epsilon$. Similarly, there is an x_3 in $2^{-2}G$ such that $\|x - x_1 - x_2 - x_3\| < 2^{-3}\epsilon$. Continuing in this way yields a sequence (x_n) such that $x_n \in 2^{-n+1}G$ and $\|x - \sum_{j=1}^{n} x_j\| < 2^{-n}\epsilon$ for each positive integer n. It follows that $p(x_n) < 2^{-n+1}$ for each n and that $x = \sum_n x_n$, and so the countable subadditivity of p implies that

$$p(x) = p\Big(\sum_n x_n\Big) \le \sum_n p(x_n) < 2.$$

Letting $s = 2$ completes the proof. ■

The guaranteed continuity of their countably subadditive seminorms is an important way in which Banach spaces differ from incomplete normed spaces, for Zabreĭko's lemma does not extend to all normed spaces. See Exercise 1.75.

1.6.4 Definition. A function f from a topological space X into a topological space Y is an *open mapping* if $f(U)$ is an open subset of Y whenever U is an open subset of X.

It is a well-known result of elementary complex analysis that every nonconstant analytic function on a connected open subset of $\mathbb{C}$ is an open mapping. Unfortunately, that fact and the unrelated one about to be proved here are both commonly called the open mapping theorem. The following result is also known as the *interior mapping principle.*

1.6.5 The Open Mapping Theorem. (J. Schauder, 1930 [208]). *Every bounded linear operator from a Banach space onto a Banach space is an open mapping.*

PROOF. Let T be a bounded linear operator from a Banach space X onto a Banach space Y. Suppose that the image under T of the open unit ball U of X is open. Let V be an open subset of X. If $x \in V$, then $x + rU \subseteq V$ for some positive r, and so $T(V)$ includes the neighborhood $Tx + rT(U)$ of Tx. It follows that $T(V)$ is open. Thus, the theorem will be proved once it is shown that $T(U)$ is open.

For each y in Y, let $p(y) = \inf\{\,\|x\| : x \in X,\, Tx = y\,\}$. If $y \in Y$ and α is a nonzero scalar, then $\{\,x : x \in X,\, Tx = \alpha y\,\} = \{\,\alpha x : x \in X,\, Tx = y\,\}$, and so

$$\begin{aligned} p(\alpha y) &= \inf\{\,\|\alpha x\| : x \in X,\, Tx = y\,\} \\ &= |\alpha| \cdot \inf\{\,\|x\| : x \in X,\, Tx = y\,\} \\ &= |\alpha|\, p(y). \end{aligned}$$

Since $p(0y) = 0 = |0|\,p(y)$ whenever $y \in Y$, it follows that $p(\alpha y) = |\alpha|\,p(y)$ for each scalar α and each y in Y. Now let $\sum_n y_n$ be a convergent series in Y. The goal is to show that $p(\sum_n y_n) \le \sum_n p(y_n)$, so $\sum_n p(y_n)$ can be assumed to be finite. Fix a positive ϵ. Let (x_n) be a sequence in X such that $Tx_n = y_n$ and $\|x_n\| < p(y_n) + 2^{-n}\epsilon$ for each n. Then $\sum_n \|x_n\| < \sum_n p(y_n) + \epsilon$, a finite number. Since X is a Banach space, the absolutely convergent series $\sum_n x_n$ converges. Now $T(\sum_n x_n) = \sum_n Tx_n = \sum_n y_n$, and so

$$p\Big(\sum_n y_n\Big) \le \Big\|\sum_n x_n\Big\| \le \sum_n \|x_n\| < \sum_n p(y_n) + \epsilon.$$

Therefore $p(\sum_n y_n) \le \sum_n p(y_n)$ since ϵ is an arbitrary positive number, and so p is countably subadditive. This also implies that $p(y_1 + y_2) \le p(y_1) + p(y_2)$ whenever $y_1, y_2 \in Y$, as can be seen by letting $y_n = 0$ when $n \ge 3$. Thus, the function p is a countably subadditive seminorm on Y, and so is continuous by Zabreĭko's lemma. Finally,

$$T(U) = \{\,y : y \in Y,\, Tx = y \text{ for some } x \text{ in } U\,\} = \{\,y : y \in Y,\, p(y) < 1\,\},$$

so $T(U)$ is open. ∎

A one-to-one map from one topological space onto another is a homeomorphism if and only if it is both continuous and open. Combining this with the preceding theorem immediately yields the following result.

1.6.6 Corollary. (S. Banach, 1929 [11]). *Every one-to-one bounded linear operator from a Banach space onto a Banach space is an isomorphism.*

That is, if a one-to-one linear operator T from a Banach space onto a Banach space is continuous, then T^{-1} is also continuous. For that reason, the preceding corollary is sometimes called the *inverse mapping theorem.*

1.6.7 Definition. Two norms on the same vector space are *equivalent* if they induce the same topology.

It follows from Corollary 1.6.6 that if $\|\cdot\|_1$ and $\|\cdot\|_2$ are two Banach norms on the same vector space X and the identity operator, viewed as a linear operator from $(X, \|\cdot\|_1)$ onto $(X, \|\cdot\|_2)$, is continuous, then a subset of X is open with respect to one of the norms if and only if it is open with respect to the other. Restating this in the language of Definition 1.6.7 gives the following result.

1.6.8 Corollary. *Suppose that $\|\cdot\|_1$ and $\|\cdot\|_2$ are two Banach norms on a vector space X and that the identity map from $(X, \|\cdot\|_1)$ to $(X, \|\cdot\|_2)$ is continuous. Then the two norms are equivalent.*

A common way of applying Corollary 1.6.8 is by finding a nonnegative number a such that $\|x\|_2 \leq a\|x\|_1$ whenever $x \in X$ and concluding that there is also some nonnegative b such that $\|x\|_1 \leq b\|x\|_2$ whenever $x \in X$, or equivalently that there is a c such that $c \geq 1$ and $c^{-1}\|x\|_1 \leq \|x\|_2 \leq c\|x\|_1$ whenever $x \in X$.

Incidentally, two Banach norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a vector space X need not be equivalent just because there is *some* one-to-one bounded linear operator that maps $(X, \|\cdot\|_1)$ onto $(X, \|\cdot\|_2)$, even if the map is an isometric isomorphism. See Exercise 1.50.

The following result is often called the *Banach-Steinhaus theorem,* since a proof of it appeared in a 1927 paper by Stefan Banach and Hugo Steinhaus [17]. The result in its full generality for Banach spaces was actually first published in 1923 by T. H. Hildebrandt [104], though special forms of it had previously appeared in a 1922 paper by Hans Hahn [98] as well as in Banach's doctoral thesis [10], also published in 1922. In particular, Hahn's proof of the result for the special case in which the family of mappings is a sequence of bounded linear functionals can easily be modified to prove the more general theorem stated here; see Exercise 1.76.

1.6.9 The Uniform Boundedness Principle. (H. Hahn, 1922 [98]; S. Banach, 1922 [10]; T. H. Hildebrandt, 1923 [104]; S. Banach and H. Steinhaus, 1927 [17]). *Let $\mathfrak{F}$ be a nonempty family of bounded linear operators from a Banach space X into a normed space Y. If $\sup\{\,\|Tx\| : T \in \mathfrak{F}\,\}$ is finite for each x in X, then $\sup\{\,\|T\| : T \in \mathfrak{F}\,\}$ is finite.*

PROOF. Let $p(x) = \sup\{\,\|Tx\| : T \in \mathfrak{F}\,\}$ for each x in X, and suppose that p is finite-valued. Notice that $p(\alpha x) = |\alpha|\,p(x)$ for each x in X and each scalar α. If $\sum_n x_n$ is a convergent series in X and $T \in \mathfrak{F}$, then

$$\left\|T\Big(\sum_n x_n\Big)\right\| = \left\|\sum_n Tx_n\right\| \leq \sum_n \|Tx_n\| \leq \sum_n p(x_n),$$

from which it follows that $p(\sum_n x_n) \leq \sum_n p(x_n)$. In particular, whenever $x_1, x_2 \in X$, letting $x_n = 0$ when $n \geq 3$ shows that $p(x_1 + x_2) \leq p(x_1) + p(x_2)$, so p is a countably subadditive seminorm on X. Therefore p is continuous, so there is some positive δ such that $p(x) \leq 1$ whenever $\|x\| \leq \delta$. It follows that $p(x) \leq \delta^{-1}$ whenever $x \in B_X$, and therefore that $\|Tx\| \leq \delta^{-1}$ whenever $T \in \mathfrak{F}$ and $x \in B_X$, that is, that $\|T\| \leq \delta^{-1}$ for each T in $\mathfrak{F}$. ■

A family $\mathfrak{F}$ of linear operators from a normed space X into a normed space Y is said to be *pointwise bounded* if, for each element x of X, the set $\{ Tx : T \in \mathfrak{F} \}$ is bounded, and is said to be *uniformly bounded* if, for each bounded subset B of X, the set $\bigcup\{ T(B) : T \in \mathfrak{F} \}$ is bounded. Notice that pointwise boundedness does not imply the boundedness of the individual operators in the family, since every family consisting of one unbounded operator is pointwise bounded. Of course, uniform boundedness does imply the boundedness of each member of the family. After only a moment's thought about the special case of an empty family of operators, it follows easily from the uniform boundedness principle that pointwise boundedness implies uniform boundedness provided each member of the family of operators is bounded and their common domain is a Banach space. This is the origin of the name of the theorem.

1.6.10 Corollary. *Let (T_n) be a sequence of bounded linear operators from a Banach space X into a normed space Y such that $\lim_n T_n x$ exists for each x in X. Define $T\colon X \to Y$ by the formula $Tx = \lim_n T_n x$. Then T is a bounded linear operator from X into Y.*

PROOF. The continuity of the vector space operations of Y and the linearity of each T_n together imply the linearity of T. Since $\sup\{ \|T_n x\| : n \in \mathbb{N} \}$ is finite for each x in X, there is a nonnegative M such that $\|T_n\| \leq M$ for each n, and so $\|T_n x\| \leq M$ for each x in B_X and each n. It follows that $\|Tx\| \leq M$ for each x in B_X, and so T is bounded. ■

A linear operator from one Banach space into another is sometimes called a *closed mapping* if it satisfies the hypotheses of the next theorem (but see Exercise 1.74). For that reason, the following result is often called the *closed mapping theorem.*

1.6.11 The Closed Graph Theorem. (S. Banach, 1932 [13]). *Let T be a linear operator from a Banach space X into a Banach space Y. Suppose that whenever a sequence (x_n) in X converges to some x in X and (Tx_n) converges to some y in Y, it follows that $y = Tx$. Then T is bounded.*

PROOF. Let $p(x) = \|Tx\|$ for each x in X. It is enough to prove that p is continuous, for then there would be a neighborhood U of 0 such that the set $p(U)$ is bounded, which would in turn imply that $T(U)$ is bounded, and that would imply the continuity of T by Theorem 1.4.2. Since p is a seminorm on X, an application of Zabreĭko's lemma will finish the proof once it is shown that p is countably subadditive. Let $\sum_n x_n$ be a convergent series in X. The proof will be finished once it is shown that $\|T(\sum_n x_n)\| \leq \sum_n \|Tx_n\|$, so it may be assumed that $\sum_n \|Tx_n\|$ is finite, which together with the completeness of Y implies that the absolutely convergent series $\sum_n Tx_n$ converges. Since $\sum_{n=1}^m x_n \to \sum_n x_n$ and $T(\sum_{n=1}^m x_n) = \sum_{n=1}^m Tx_n \to \sum_n Tx_n$ as $m \to \infty$, it follows from the hypotheses of the theorem that $\sum_n Tx_n = T(\sum_n x_n)$. Therefore

$$\left\|T\Big(\sum_n x_n\Big)\right\| = \left\|\sum_n Tx_n\right\| \leq \sum_n \|Tx_n\|,$$

which shows that p is countably subadditive and finishes the proof. ■

It is customary to derive all three of the major theorems of this section directly or indirectly from the Baire category theorem. This is exactly what has been done here, since Theorem 1.3.14, a weak version of the Baire category theorem, was used in the proof of Zabreĭko's lemma. See Exercise 1.76 for a different derivation of the uniform boundedness principle that does not use any form of the Baire category theorem. The argument, essentially due to Hahn, is of a type called a *gliding hump* argument. The only use of completeness in the argument is to assure that a certain absolutely convergent series converges.

Exercises

1.64 Suppose that X is a Banach space and $T\colon X \to \ell_\infty$ is a linear operator. For each n in $\mathbb{N}$, let $(Tx)_n$ be the n^{th} term of Tx and let f_n be the linear functional on X that maps x to $(Tx)_n$. Prove that T is bounded if and only if each f_n is bounded.

1.65 Repeat the preceding exercise, replacing ℓ_∞ by ℓ_1.

1.66 Prove that if T is a bounded linear operator from a Banach space X onto a Banach space Y, then whenever a sequence (y_n) converges to a limit y in Y, there is a sequence (x_n) converging to an x in X such that $Tx = y$ and $Tx_n = y_n$ for each n.

1.67 Suppose that X and Y are metric spaces and that f is a function from X into Y. Prove that the graph $\{\,(x, f(x)) : x \in X\,\}$ of f is closed in $X \times Y$ if and only if $y = f(x)$ whenever a sequence (x_n) converges to x in X and $(f(x_n))$ converges to y in Y. (This is the source of the name of the closed graph theorem. Suppose that X and Y are Banach spaces and that T is a linear operator from X into Y. Then the closed graph theorem says that T is bounded if its graph is closed in $X \times Y$. In fact, the operator T is bounded if *and only if* its graph is closed in $X \times Y$. See the next exercise.)

1.68 Prove the following converse of the closed graph theorem: If f is a continuous function from a topological space X into a Hausdorff space Y, then the graph of f is closed in $X \times Y$.

1.69 Let X be a normed space and let $X^* = B(X, \mathbb{F})$. Suppose that (x_n) is a sequence in X such that $\sum_n x^* x_n$ converges whenever $x^* \in X^*$. Show that the mapping $x^* \mapsto \sum_n x^* x_n$ is a bounded linear functional on X^*.

1.70 Suppose that $1 \le p \le \infty$ and that T is a linear operator from $L_p[0,1]$ into itself with the property that if (f_n) is a sequence in $L_p[0,1]$ that converges almost everywhere to some f in $L_p[0,1]$, then (Tf_n) converges almost everywhere to Tf. Prove that T is bounded.

1.71 Suppose that X and Y are Banach spaces and that $\mathfrak{F}$ is a family of continuous functions from Y into a Hausdorff space such that whenever $y_1, y_2 \in Y$ and $y_1 \neq y_2$, there is an f_{y_1,y_2} in $\mathfrak{F}$ for which $f_{y_1,y_2}(y_1) \neq f_{y_1,y_2}(y_2)$. Prove that if T is a linear operator from X into Y such that $f \circ T$ is continuous for each f in $\mathfrak{F}$, then T is bounded.

1.72 Derive Corollary 1.6.6 from the closed graph theorem.

1.73 Obtain the uniform boundedness principle directly from Theorem 1.3.14. (With all notation as in the statement of the uniform boundedness principle, let $C = \{\, x : x \in X,\ \|Tx\| \le 1 \text{ for each } T \text{ in } \mathfrak{F} \,\}$.)

1.74 The term *closed mapping* is used both for a linear operator between Banach spaces that satisfies the hypotheses of the closed graph theorem and for a mapping from one topological space to another that always maps closed sets onto closed sets. The two meanings are not equivalent for linear operators between Banach spaces, as will be shown in this exercise.

(a) Suppose that X and Y are normed spaces and T is a linear operator from X into Y that is neither one-to-one nor the zero operator. Find a closed subset F of X such that $T(F)$ is not closed in Y.

(b) Find a linear operator T that satisfies the hypotheses of the closed graph theorem even though there is a closed subset F of the domain of T such that $T(F)$ is not closed in the range of T.

1.75 The results of this section cannot in general be extended to incomplete normed spaces. To see this, let X be the vector space of finitely nonzero sequences equipped with the ℓ_1 norm. Let Y have the same underlying vector space as X, but equipped with the ℓ_∞ norm.

(a) Let $f_m((\alpha_n)) = m \cdot \alpha_m$ for each element (α_n) of X and each positive integer m. Use the family $\{\, f_m : m \in \mathbb{N} \,\}$ to show that the uniform boundedness principle does not extend to incomplete normed spaces.

(b) Show that the "identity" operator from X onto Y is a bounded linear mapping that is not open.

(c) Use the "identity" operator from Y onto X to show that the closed graph theorem does not extend to incomplete normed spaces.

(d) Show that the ℓ_1 norm is a countably subadditive seminorm on Y that is not continuous.

1.76 The purpose of this exercise is to produce a proof of the uniform boundedness principle without using Theorem 1.3.14 or any other form of the Baire category theorem. To this end, suppose that $\mathfrak{F}$ is a nonempty collection of bounded linear operators from a Banach space X into a normed space Y such that $\sup\{\, \|T\| : T \in \mathfrak{F}\,\} = +\infty$. The goal is to find an x in X such that $\sup\{\, \|Tx\| : T \in \mathfrak{F}\,\} = +\infty$.

(a) The proof is based on the existence of sequences (T_n) and (x_n) in $\mathfrak{F}$ and X respectively such that the following two conditions are satisfied for each positive integer n:

$$\|T_n x_n\| \geq n + \sum_{j=1}^{n-1} \|T_n x_j\| \quad (\text{or} \geq 1 \text{ if } n = 1);$$
$$\|x_n\| \leq 2^{-n} \min\{\, \|T_j\|^{-1} : j < n \,\} \quad (\text{or} \leq 2^{-1} \text{ if } n = 1).$$

Argue that the only obstacle to the inductive construction of the sequences is the existence of an x with the desired property. The existence of such a pair of sequences may therefore be assumed.

(b) Show that the series $\sum_n x_n$ converges to some x in X.

(c) Show that $\sum_{j=n+1}^{\infty} \|T_n x_j\| \leq 1$ for each n.

(d) Show that $\|T_n x\| \geq n - 1$ for each n, so $\sup\{\, \|Tx\| : T \in \mathfrak{F}\,\} = +\infty$.

This proof is essentially from Hahn's 1922 paper [98, pp. 6–8], though he stated the result only for sequences of linear functionals. This is called a *gliding hump* argument. The sequences (T_n) and (x_n) are chosen so that $\sum_n x_n$ converges to some x in X and, for each n, the major contribution to $T_n x$ comes from $T_n x_n$. If $\big(\|T_n x_j\|\big)_{j=1}^{\infty}$ is considered to be a sequence that depends on the parameter n, then the sequence has a "hump" in it at the n^{th} term. As n increases, this hump glides forward and has unbounded height. (Incidentally, the argument can be simplified very slightly by replacing $\sum_{j=1}^{n-1} \|T_n x_j\|$ by $\|T_n(\sum_{j=1}^{n-1} x_j)\|$ in (a), but then it is not obvious that $\|T_n x_n\|$ must be the dominant term of the sequence $\big(\|T_n x_j\|\big)_{j=1}^{\infty}$, and therefore it is harder to see the hump.) Gliding hump arguments of this form probably first appeared in work by Henri Lebesgue from 1905; see [149] and [150, pp. 86–88]. Hahn specifically stated in his paper that the basic method for his proof was taken from a 1909 paper by Lebesgue [151, p. 61]. See [64, pp. 138–142] for more on the history of gliding hump arguments.

1.7 Quotient Spaces

Recall the following definition from linear algebra.

1.7.1 Definition. Let M be a subspace of a vector space X. The *quotient space* or *factor space* X/M (read "X modulo M" or "X mod M") is the vector space whose underlying set is the collection $\{\, x + M : x \in X \,\}$ of all translates of M along with the vector space operations given by the formulas

$$(x + M) + (y + M) = (x + y) + M$$

and

$$\alpha \cdot (x + M) = (\alpha \cdot x) + M.$$

For each x in X, the translate $x + M$ is called the *coset of M containing x.*

With M and X as in this definition, it follows that two elements $x + M$ and $x' + M$ of X/M are equal if and only if $x - x' \in M$. Easy arguments based on this show that the addition of elements and multiplication of elements by scalars given in the definition are well-defined, that is, that if $\alpha \in \mathbb{F}$ and x, x', y, y' are elements of X such that $x + M = x' + M$ and $y + M = y' + M$, then $(x + M) + (y + M) = (x' + M) + (y' + M)$ and $\alpha(x + M) = \alpha(x' + M)$. The fact that X/M with these operations is a vector space then follows by routine verifications. Notice that the zero of X/M is M, for which $0 + M$ is usually written in this context, and that $-(x + M) = (-x) + M$ whenever $x \in X$.

Since two cosets $x + M$ and $x' + M$ are equal if and only if $x' - x \in M$, and therefore if and only if $x' \in x + M$, it follows that $x + M$ and $x' + M$ are either equal or disjoint. Thus, the sets that are the elements of X/M partition X into equivalence classes, with two elements of X lying in the same equivalence class if and only if their difference lies in M. For this reason, the elements of X/M are sometimes treated as if they were just those of X, except that two elements are considered to be the same when they differ only by an element of M. Here is one familiar example of this practice.

1.7.2 Example. Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω. Let X be the vector space of all μ-measurable scalar-valued functions on Ω that are μ-integrable, and let M be the subspace of X consisting of all members of X that are zero almost everywhere. Then the vector space underlying $L_1(\Omega, \Sigma, \mu)$ is actually X/M, though in practice it is often treated as if it were X. Similar remarks apply to the spaces $L_p(\Omega, \Sigma, \mu)$ such that $1 < p \leq \infty$.

The notation of Definition 1.7.1 suggests that the sum of two cosets is just the algebraic sum of the two sets as defined in Section 1.1. It is easy to verify that this is indeed so. It is also easy to check that the product of a scalar α with a coset $x + M$ is just the product of α with the set $x + M$ as such products were defined in Section 1.1 *provided α is nonzero.* There is a trap waiting when $\alpha = 0$, since

$$0(x + M) = \begin{cases} 0 + M = M & \text{in the sense of Definition 1.7.1;} \\ \{0\} & \text{in the sense of Section 1.1.} \end{cases}$$

It is thus all right to treat the vector space operations of Definition 1.7.1 as being the set operations of Section 1.1 as long as this one discrepancy is kept in mind. The context should always prevent confusion about the meaning of $0(x+M)$.

Now suppose that M is a subspace of a normed space X. It is reasonable to ask if the norm of X induces a norm on X/M in some natural way. Such a norm clearly cannot come from the "formula" $\|x+M\| = \|x\|$ when $M \neq \{0\}$, for $y+M = 0+M$ whenever $y \in M$ even if $\|y\| \neq \|0\|$. This is one situation in which it is helpful to think first about distance and then recover the norm from the notion of distance. There is a natural way to define the distance between two cosets $x+M$ and $y+M$, namely, the way in which the distance between subsets of a metric space is defined in Section 1.1:

$$d(x+M, y+M) = \inf\{\, \|v-w\| : v \in x+M,\ w \in y+M \,\}. \tag{1.1}$$

Also, the formula from Section 1.1 for the distance between a point and a set in a metric space can be used to find the distance between an element x of X and a member $y+M$ of X/M:

$$d(x, y+M) = \inf\{\, \|x-w\| : w \in y+M \,\}.$$

It follows that $d(x, y+M) = d(x+M, y+M)$ whenever $x, y \in X$, since

$$\begin{aligned}\{\, v-w : v \in x+M,\ w \in y+M \,\} &= \{\, (x+z_1)-(y+z_2) : z_1, z_2 \in M \,\}\\ &= \{\, x-(y+z_2-z_1) : z_1, z_2 \in M \,\}\\ &= \{\, x-(y+z) : z \in M \,\}\\ &= \{\, x-w : w \in y+M \,\}.\end{aligned}$$

Notice that if $x \in \overline{M} \setminus M$, then $0 \le d(x+M, 0+M) = d(x, M) = 0$, and therefore $d(x+M, 0+M) = 0$ even though $x+M \neq 0+M$. It follows that if the function d of (1.1) is to have any hope of being a metric on X/M, then the set $\overline{M} \setminus M$ must be empty; that is, the set M must be closed. Therefore, for the rest of this paragraph it will be assumed that M is a *closed* subspace of X. If the formula from (1.1) does now define a metric on X/M (which it does, as will be shown in the comments following Theorem 1.7.4), and if this metric is induced by a norm, then that norm must measure the distance from a coset to the origin of X/M.

1.7.3 Definition. Let M be a closed subspace of a normed space X. The *quotient norm* of X/M is given by the formula $\|x+M\| = d(x+M, 0+M)$.

With M and X as in the preceding definition, the quotient norm of a coset $x+M$ can also be interpreted to be the distance from the point x to the set M, or as the distance from the origin of X to the set $x + M$, since $d(x, M) = d(x + M, 0 + M) = d(0, x + M)$. Therefore

$$\|x + M\| = \inf\{\, \|x - z\| : z \in M \,\} = \inf\{\, \|x + z\| : z \in M \,\}$$

whenever $x \in X$.

1.7.4 Theorem. *If M is a closed subspace of a normed space X, then the quotient norm of X/M is a norm.*

PROOF. Suppose that $x, y \in X$ and that α is a scalar. Since M is closed, it follows that $d(x, M) = 0$ if and only if $x \in M$, that is, that $\|x + M\| = 0$ if and only if $x + M = 0 + M$. If $\alpha \neq 0$, then

$$\|\alpha(x + M)\| = d(\alpha x, M) = d(\alpha x, \alpha M) = |\alpha|\, d(x, M) = |\alpha|\, \|x + M\|,$$

and

$$\|0(x + M)\| = \|0 + M\| = 0 = |0|\, \|x + M\|.$$

To verify the triangle inequality, first observe that whenever $z_1, z_2 \in M$,

$$\begin{aligned}\|(x + M) + (y + M)\| &= \|(x + y) + M\| \\ &\leq \|x + y + z_1 + z_2\| \\ &\leq \|x + z_1\| + \|y + z_2\|.\end{aligned}$$

Taking appropriate infima shows that

$$\|(x + M) + (y + M)\| \leq \|x + M\| + \|y + M\|.$$

Thus, the quotient norm has all the properties required of a norm. ■

When X and M are as in the preceding theorem, the metric induced by the quotient norm of X/M is given by the formula from (1.1), since

$$\begin{aligned}d(x + M, y + M) &= d\big((x - y) + M, 0 + M\big) \\ &= \|(x - y) + M\| \\ &= \|(x + M) - (y + M)\|\end{aligned}$$

whenever $x, y \in X$.

Henceforth, whenever a quotient space of a normed space by one of its closed subspaces is treated as a normed space, a metric space, or a topological space though no norm, metric, or topology has been specified, it is the quotient norm, the metric induced by that norm, or the topology induced by that metric that is implied.

1.7.5 Example. In the Euclidean space $\mathbb{R}^2$, let $M = \{\, (\alpha, \beta) : \alpha = 0 \,\}$. Then $\mathbb{R}^2/M$ is the collection of all vertical lines in the plane, with the norm of each such line being its distance from the origin, that is, the absolute value of its α-intercept. Some of the proofs in this section are based on geometric ideas that are easier to visualize if this example is kept in mind.

1.7.6 Proposition. *Let M be a closed subspace of a normed space X.*

(a) *If $x \in X$, then $\|x\| \ge \|x+M\|$.*

(b) *If $x \in X$ and $\epsilon > 0$, then there is an x' in X such that $x'+M = x+M$ and $\|x'\| < \|x+M\| + \epsilon$.*

PROOF. Since $\|x-0\| \ge d(x,M) = \|x+M\|$ for each x in X, part (a) holds. For (b), suppose that $x \in X$ and $\epsilon > 0$. Let y be an element of M such that $\|x-y\| < d(x,M)+\epsilon = \|x+M\|+\epsilon$. Then $x-y$ is the desired x'. ■

With X and M as in the preceding proposition, suppose that x and y are elements of X such that $\|(x-y)+M\| < \delta$ for some positive δ. By part (b) of the proposition, there is a y' in X such that $(x-y')+M = (x-y)+M$ and $\|x-y'\| < \delta$. This observation is useful in the proof of the next theorem.

1.7.7 Theorem. *If M is a closed subspace of a Banach space X, then X/M is also a Banach space.*

PROOF. Suppose that (x_n+M) is a Cauchy sequence in X/M. It is enough to prove that some subsequence of (x_n+M) has a limit, for then the entire sequence will converge to the same limit. By thinning, it may be assumed that $\|(x_n - x_{n+1}) + M\| < 2^{-n}$ for each n. By the remark preceding the theorem, there is some x_2' such that $(x_1 - x_2') + M = (x_1 - x_2) + M$ and $\|x_1 - x_2'\| < 2^{-1}$. Since $x_2' + M = x_2 + M$, it may be assumed that $x_2' = x_2$. There is some x_3' such that $(x_2 - x_3') + M = (x_2 - x_3) + M$ and $\|x_2 - x_3'\| < 2^{-2}$. It may be assumed that $x_3' = x_3$. By an obvious induction argument, it may be assumed that $\|x_n - x_{n+1}\| < 2^{-n}$ for each n. Then the Cauchy sequence (x_n) converges to some x in X. For each n,

$$\|(x_n + M) - (x + M)\| = \|(x_n - x) + M\| \le \|x_n - x\|,$$

and so $x_n + M \to x + M$ as $n \to \infty$. ■

Suppose now that M is a closed subspace of a normed space X not known to be complete. It is natural to ask whether the completeness of X/M would imply that of X. In general, the answer is no. A trivial counterexample is obtained by letting X be any incomplete normed space, letting $M = X$, and noting that X/M is complete even though X is not. It turns out, though, that X would have to be complete if both X/M and M were so.

1.7.8 Definition. Let P be a property defined for normed spaces. Suppose that whenever X is a normed space with a closed subspace M such that two of the spaces X, M, and X/M have the property, then the third must also have it. Then P is a *three-space* property.

1.7.9 Theorem. *Completeness is a three-space property.*

PROOF. Let M be a closed subspace of a normed space X. If X is complete, then so are M and X/M. Therefore all that needs to be checked is that X is complete if both M and X/M are.

Suppose that M and X/M are Banach spaces. Let (x_n) be a Cauchy sequence in X. Since $\|(x_n - x_m) + M\| \leq \|x_n - x_m\|$ whenever $m, n \in \mathbb{N}$, the sequence $(x_n + M)$ is Cauchy in X/M and so converges to some $y + M$. By Proposition 1.7.6 (b), there is for each positive integer n some y_n in X such that $y_n + M = (x_n - y) + M$ and $\|y_n\| < \|(x_n - y) + M\| + 2^{-n}$. Then $\lim_n y_n = 0$, so $(x_n - y_n - y)$ is a Cauchy sequence in M and therefore has a limit z in M. It follows that $x_n = (x_n - y_n - y) + y_n + y \to z + y$ as $n \to \infty$, so X is complete. ■

When a set X has been partitioned into a collection Y of equivalence classes, the function from X onto Y that maps each member of X to its equivalence class is often called the *projection* from X onto Y. If X is a normed space and the equivalence classes are the members of some quotient space X/M such that M is a closed subspace of X, then it is customary to call the projection from X onto X/M a *quotient map,* partly for the obvious reason but also because of a relationship to topological quotient maps that will be mentioned later.

1.7.10 Definition. Let M be a closed subspace of a normed space X. Then the *quotient map* from X onto X/M is the function π defined by the formula $\pi(x) = x + M$.

1.7.11 Lemma. *If M is a closed subspace of a normed space X and π is the quotient map from X onto X/M, then the image under π of the open unit ball of X is the open unit ball of X/M.*

PROOF. This proof uses both parts of Proposition 1.7.6. Let U_X and $U_{X/M}$ be the open unit balls of X and X/M respectively. If $x \in U_X$, then $\|\pi(x)\| = \|x + M\| \leq \|x\| < 1$, so $\pi(U_X) \subseteq U_{X/M}$. If $y + M \in U_{X/M}$, then there is a y' in U_X such that $\pi(y') = y' + M = y + M$. Therefore $U_{X/M} \subseteq \pi(U_X)$, so $\pi(U_X) = U_{X/M}$. ■

1.7.12 Proposition. *Let M be a closed subspace of a normed space X. Then the quotient map π from X onto X/M is a bounded linear operator that is also an open mapping and has M as its kernel. If $M \neq X$, then $\|\pi\| = 1$.*

PROOF. The linearity of π follows immediately from the definition of the vector space operations of X/M, and the fact that $\ker(\pi) = M$ is almost as immediate. By the lemma, the linear operator π maps the open unit ball of X onto a bounded subset of X/M, which implies that π is bounded. It is also clear from the lemma that $\|\pi\| = 1$ if $X/M \neq \{M\}$, that is, if $M \neq X$.

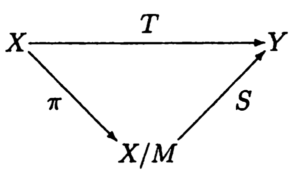
Finally, let x be an element of an open subset V of X. To show that π is an open mapping, it is enough to find a neighborhood W of $\pi(x)$ such that $W \subseteq \pi(V)$. Let U_X be the open unit ball of X. Then $x + rU_X \subseteq V$ for some positive r, so $\pi(x) + r\pi(U_X) \subseteq \pi(V)$. Since $\pi(U_X)$ is the open unit ball of X/M, the neighborhood $\pi(x) + r\pi(U_X)$ of $\pi(x)$ is an acceptable W. ■

Recall that a *topological quotient map* is a function f from a topological space X onto a topological space Y such that a subset A of Y is open in Y if and only if $f^{-1}(A)$ is open in X, and that if there is such a map f from X onto Y, then the topology of Y is called the *quotient topology* induced by f. Of course, a function from one topological space onto another that is both continuous and an open mapping is a topological quotient map. It therefore follows from Proposition 1.7.12 that the quotient map π from a normed space X onto X/M, where M is a closed subspace of X, is a topological quotient map, and the topology of X/M is the quotient topology induced by π.

Another important class of topological quotient maps consists of the bounded linear operators from one Banach space onto another. Each such map actually is a topological quotient map since it is both continuous and open. More generally, let T be a bounded linear operator from a Banach space X into a Banach space Y such that the range of T is closed in Y. Then T, viewed as an operator onto the Banach space $T(X)$, is a topological quotient map, so the norm topology that $T(X)$ inherits from Y is also the quotient topology induced by T.

An important feature of a closed-range bounded linear operator from one Banach space into another is that its range is an isomorphic copy of a certain quotient space of its domain. The next theorem is the main ingredient needed to prove this.

1.7.13 Theorem. *Suppose that X and Y are normed spaces and that T is a linear operator from X into Y, not assumed to be bounded. Suppose further that M is a closed subspace of X such that $M \subseteq \ker(T)$ and that π is the quotient map from X onto X/M. Then there is a unique function S from X/M into Y such that $T = S \circ \pi$, that is, such that the following diagram commutes.*



This map S is linear and has the same range as T. The operator S is an open mapping if and only if T is an open mapping, and is bounded if and only if T is bounded. If T is bounded, then $\|S\| = \|T\|$.

PROOF. Let $S(x+M) = Tx$ whenever $x \in X$. If $x + M = x' + M$, then $x - x' \in M \subseteq \ker(T)$, and so $Tx = Tx'$. Thus, there is no ambiguity in the definition of S. The linearity of S now follows immediately from that of T. Notice that $T = S \circ \pi$, and that if $S'\colon X/M \to Y$ is such that $T = S' \circ \pi$, then $S'(x+M) = S'\big(\pi(x)\big) = Tx = S(x+M)$ for each member $x + M$ of X/M. This proves the claimed existence and uniqueness of S. Also,

$$\{\, S(x+M) : x+M \in X/M \,\} = \{\, S\big(\pi(x)\big) : x \in X \,\} = \{\, Tx : x \in X \,\},$$

so the ranges of S and T are the same.

If U_X and $U_{X/M}$ are the open unit balls of X and X/M respectively, then $\pi(U_X) = U_{X/M}$ by Lemma 1.7.11, and so

$$\begin{aligned} \sup\{\, \|S(x+M)\| : x+M \in U_{X/M} \,\} &= \sup\{\, \|S\big(\pi(x)\big)\| : x \in U_X \,\} \\ &= \sup\{\, \|Tx\| : x \in U_X \,\}. \end{aligned}$$

It follows that S is bounded if and only if T is bounded, and that if S and T are bounded then the common value of the above suprema is the norm of each.

Since the composite of two open mappings is an open mapping, it follows that T is an open mapping whenever S is. Conversely, suppose that T is an open mapping. If U is an open subset of X/M, then

$$S(U) = S\Big(\pi\big(\pi^{-1}(U)\big)\Big) = T\big(\pi^{-1}(U)\big),$$

an open set. This shows that S is an open mapping and completes the proof. ■

The preceding theorem may have a familiar ring to it, for it is the normed-space analog of an important result from abstract algebra: If N is a normal subgroup of a group G and f is a group homomorphism from G into a group H such that $N \subseteq \ker(f)$, then there is a unique group homomorphism f_0 from the quotient group G/N into H such that $f(g) = f_0(gN)$ for each g in G. This result is used to prove the first isomorphism theorem for groups, which says that if $f\colon G \to H$ is a group homomorphism, then $G/\ker(f)$ and $f(G)$ are isomorphic as groups. Theorem 1.7.13 can be used to obtain an analogous result for Banach spaces.

1.7.14 The First Isomorphism Theorem for Banach Spaces. *Suppose that X and Y are Banach spaces and that $T \in B(X,Y)$. Suppose further that the range of T is closed in Y. Then $X/\ker(T) \cong T(X)$.*

PROOF. As the inverse image under T of the closed subspace $\{0\}$ of Y, the kernel of T is a closed subspace of X. Let $S\colon X/\ker(T) \to Y$ be the linear operator obtained from Theorem 1.7.13 by letting $M = \ker(T)$. Since

$T = S\pi$,

$$\begin{aligned}\ker(S) &= \{\, x + \ker(T) : x \in X,\ S(\pi x) = 0 \,\} \\ &= \{\, x + \ker(T) : x \in \ker(T) \,\} \\ &= \{0 + \ker(T)\}.\end{aligned}$$

It follows that S is a one-to-one bounded linear operator from the Banach space $X/\ker(T)$ onto the Banach space $T(X)$, and therefore is an isomorphism by Corollary 1.6.6. ∎

This section concludes with two applications of quotient maps. The first gives a nice test for the continuity of a finite-rank linear operator between normed spaces.

1.7.15 Theorem. *Suppose that T is a finite-rank linear operator from a normed space X into a normed space Y. Then T is bounded if and only if its kernel is a closed subset of X.*

PROOF. If T is bounded, then the kernel of T is closed since it is the inverse image under T of the closed subset $\{0\}$ of Y. Conversely, suppose that $\ker(T)$ is closed. Let $M = \ker(T)$ and let $S\colon X/M \to Y$ be as in Theorem 1.7.13. Then $S\big(x + \ker(T)\big) = 0$ exactly when $Tx = 0$, that is, exactly when $x \in \ker(T)$. Thus, the kernel of S is the one-element set $\{\ker(T)\}$, and so S is one-to-one. Since $T(X)$, the range of S, is finite-dimensional, the domain of the one-to-one linear operator S is also finite-dimensional, and so S is bounded. By Theorem 1.7.13, the operator T must also be bounded. ∎

Incidentally, a much stronger statement can be made for nonzero linear functionals. If f is a nonzero linear functional on a normed space X, then the kernel of f is either closed in X or dense in X, depending on whether or not f is bounded.

1.7.16 Proposition. *Suppose that f is an unbounded linear functional on a normed space X. Then the kernel of f is dense in X.*

PROOF. If α is a nonzero element of a balanced subset B of $\mathbb{F}$ and β is a scalar such that $|\beta| \le |\alpha|$, then $|\alpha^{-1}\beta| \le 1$, and so $\beta = \alpha^{-1}\beta\alpha \in B$. It follows that the only balanced unbounded subset of $\mathbb{F}$ is $\mathbb{F}$ itself. Since the open unit ball U_X of X is balanced and f is an unbounded linear functional, the set $f(U_X)$ is a balanced unbounded subset of $\mathbb{F}$ and so equals $\mathbb{F}$. If V is an open ball in X, then $V = x + rU_X$ for some x in X and some positive r, so $f(V) = f(x + rU_X) = f(x) + rf(U_X) = \mathbb{F}$, which implies that f takes on the value zero somewhere on V. Thus, the kernel of f intersects every open ball in X, and so is dense in X. ∎

For the second application, suppose that M and N are closed subspaces of a normed space X. It is easy to see that $M+N$ is also a subspace of X. However, the set $M+N$ might not be closed; see Exercise 1.84. There is one situation, however, in which it must be.

1.7.17 Proposition. *Let M and N be closed subspaces of a normed space X. If either M or N is finite-dimensional, then $M+N$ is a closed subspace of X.*

PROOF. Without loss of generality, assume that M is finite-dimensional. Let π be the quotient map from X onto X/N. By Corollary 1.4.20, the finite-dimensional subspace $\pi(M)$ of X/N is closed, so $\pi^{-1}(\pi(M))$ is a closed subspace of X. It is easy to check that $M+N=\pi^{-1}(\pi(M))$. ∎

Exercises

1.77 Prove that $\|(\alpha_n)+c_0\| = \limsup_n|\alpha_n|$ for each element $(\alpha_n)+c_0$ of ℓ_∞/c_0.

1.78 Let $M = \{\, f : f \in C[0,1],\ f(0) = 0 \,\}$, a closed subspace of $C[0,1]$. Find a simple expression for the quotient norm of $C[0,1]/M$. What familiar normed space is isometrically isomorphic to this quotient space?

1.79 Give an example to show that the isomorphism guaranteed by the first isomorphism theorem need not be an isometric isomorphism. Exercise 1.49 might help.

1.80 Display a bounded linear operator from a Banach space onto an incomplete normed space. What does this say about trying to extend the first isomorphism theorem to incomplete normed spaces?

1.81 Suppose that T is a bounded finite-rank linear operator from a normed space X into a normed space Y. Prove that $X/\ker(T) \cong T(X)$, whether or not either X or Y is complete.

1.82 Show that the conclusion of Theorem 1.7.15 does not always hold for linear operators not having finite rank. Exercise 1.75 (c) may be helpful.

1.83 Prove that finite-dimensionality is a three-space property.

1.84 The purpose of this exercise is to show that the sum of two closed subspaces of a Banach space need not be closed. Let M and N be the closed subspaces of c_0 defined by the formulas

$$M = \{\, (\alpha_n) : (\alpha_n) \in c_0,\ \alpha_m = m\alpha_{m-1} \text{ for each even } m \,\};$$
$$N = \{\, (\alpha_n) : (\alpha_n) \in c_0,\ \alpha_m = 0 \text{ for each odd } m \,\}.$$

Show that $M+N$ is a proper dense subspace of c_0.

1.85 In c_0, let $M_1 = \{\, (\alpha_n) : \alpha_1 = 0 \,\}$ and $M_2 = \{\, (\alpha_n) : \alpha_1 = \alpha_2 = 0 \,\}$. Show that c_0/M_1 is not isomorphic to c_0/M_2, even though M_1 is isometrically isomorphic to M_2.

1.86 The purpose of this exercise is to prove this theorem: *Suppose that X and Y are normed spaces and that M is a closed subspace of X such that X/M is finite-dimensional. If $T \in B(M,Y)$, then there is a T_0 in $B(X,Y)$ such that T_0 agrees with T on M. That is, every bounded linear operator with domain M has a bounded linear extension to X.*

(a) Show that the theorem is true if the dimension of X/M is zero. Now assume that the dimension of X/M is a positive integer n. Let $x_1 + M, \ldots, x_n + M$ be a basis for X/M and let $Z = \langle x_1, \ldots, x_n \rangle$. Show that for each x in X there is a unique $m(x)$ in M and a unique $z(x)$ in Z such that $x = m(x) + z(x)$.

(b) Show that the mappings $x \mapsto m(x)$ and $x \mapsto z(x)$ are bounded linear operators from X onto M and Z respectively.

(c) Prove the theorem.

1.87 Find the second and third isomorphism theorems for groups in your favorite abstract algebra text; for example, see [106].

(a) State and prove a Banach-space analog of the second isomorphism theorem. You might need to impose some conditions on the pertinent subspaces beyond just requiring them to be closed.

(b) Do the same for the third isomorphism theorem.

1.8 Direct Sums

Recall that the *vector space sum* of the vector spaces $X_1, \ldots, X_n$ in a nonempty finite ordered list is the vector space whose underlying set is the Cartesian product $X_1 \times \cdots \times X_n$ and which has the vector space operations given by the formulas

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n);$$
$$\alpha \cdot (x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n).$$

If $X_1, \ldots, X_n$ are normed spaces, then there is a way to norm their vector space sum that is suggested by the norm of Euclidean n-space.

1.8.1 Definition. Let $X_1, \ldots, X_n$ be normed spaces with respective norms $\|\cdot\|_{X_1}, \ldots, \|\cdot\|_{X_n}$. The *(external) direct sum* or *direct product* of $X_1, \ldots, X_n$ is the normed space whose underlying vector space is the vector space sum of $X_1, \ldots, X_n$ and whose norm is the *direct sum norm* given by the formula

$$\|(x_1, \ldots, x_n)\| = \left(\sum_{j=1}^{n} \|x_j\|_{X_j}^2 \right)^{1/2}.$$

This normed space is denoted by $X_1 \oplus \cdots \oplus X_n$.

The subscripts that appear on the norms of $X_1, \dots, X_n$ in the preceding definition were included only for clarity, and will usually be omitted when there is no possibility of confusion.

Of course, it must be shown that the formula in Definition 1.8.1 actually does define a norm on the vector space sum of $X_1, \dots, X_n$. The triangle inequality follows from the triangle inequality for the norm of Euclidean n-space, for if $(x_1, \dots, x_n), (y_1, \dots, y_n) \in X_1 \times \cdots \times X_n$, then

$$\begin{aligned}\|(x_1, \dots, x_n) + (y_1, \dots, y_n)\| &= \Big(\sum_{j=1}^n \|x_j + y_j\|^2\Big)^{1/2} \\ &\le \Big(\sum_{j=1}^n (\|x_j\| + \|y_j\|)^2\Big)^{1/2} \\ &\le \Big(\sum_{j=1}^n \|x_j\|^2\Big)^{1/2} + \Big(\sum_{j=1}^n \|y_j\|^2\Big)^{1/2} \\ &= \|(x_1, \dots, x_n)\| + \|(y_1, \dots, y_n)\|.\end{aligned}$$

It is easy to check that the direct sum norm has the other properties required of a norm.

It should be mentioned that there is no universal agreement on the best way to define the direct sum norm. For example, the norms defined by the formulas

$$\|(x_1, \dots, x_n)\|_1 = \sum_{j=1}^n \|x_j\|$$

and

$$\|(x_1, \dots, x_n)\|_\infty = \max\{\|x_1\|, \dots, \|x_n\|\}$$

are often used instead of the norm given in Definition 1.8.1. Fortunately, these three norms turn out to be equivalent; see Exercise 1.88. The reason for choosing the particular norm used here is made clear in Section 1.10.

Just as the symbol $\sum_{j=1}^n \alpha_j$ is often used to represent the sum of scalars $\alpha_1, \dots, \alpha_n$, there are compact "sigma" notations to represent the direct sum of normed spaces $X_1, \dots, X_n$. Unfortunately, there seem to be about as many such notations in use as there are people to invent them. For example, the notations $\sum_{j=1}^n \oplus_{\ell_2} X_j$, $\bigoplus_{j=1}^n {}_{\ell_2} X_j$, and $\big(\sum_{j=1}^n X_j\big)_{\ell_2}$ are all used, where the ℓ_2 indicates in an obvious way the nature of the direct sum norm and is sometimes omitted or replaced by ℓ_2^n or just 2.

When a normed space direct sum has only one summand X_1, the notation of Definition 1.8.1 does not distinguish between the summand X_1 and the direct sum X_1. This is no problem, for in this situation it is common practice to identify the direct sum with the summand. Since the map

$x \mapsto (x)$ from the summand onto the direct sum is clearly an isometric isomorphism, the convention is justified.

When $(X_1, d_1), \ldots, (X_n, d_n)$ are metric spaces, there is a standard way to define a *product metric* on $X_1 \times \cdots \times X_n$, namely, by the formula

$$d_p((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \left(\sum_{j=1}^{n} (d_j(x_j, y_j))^2 \right)^{1/2};$$

see Definition B.50 in Appendix B. Furthermore, this metric induces the usual product topology on $X_1 \times \cdots \times X_n$. When $X_1, \ldots, X_n$ are normed spaces, the metric induced on $X_1 \times \cdots \times X_n$ by the direct sum norm is precisely the product metric d_p, which immediately yields the following result.

1.8.2 Theorem. *Let $X_1, \ldots, X_n$ be normed spaces. Then the product metric induced on $X_1 \times \cdots \times X_n$ by the metrics of $X_1, \ldots, X_n$ is the same as the metric induced by the direct sum norm, so the product topology of $X_1 \times \cdots \times X_n$ is the same as the topology induced by the direct sum norm.*

The next result says that the direct sum of normed spaces $X_1, \ldots, X_n$ always includes isometrically isomorphic copies of the X_j's in the places where one would most expect to find them.

1.8.3 Proposition. *Let $X_1, \ldots, X_n$ be normed spaces. For each integer j such that $1 \leq j \leq n$, let*

$$X_j' = \{ (x_1, \ldots, x_n) : (x_1, \ldots, x_n) \in X_1 \oplus \cdots \oplus X_n,\ x_k = 0 \text{ when } k \neq j \}.$$

Then each X_j' is a closed subspace of $X_1 \oplus \cdots \oplus X_n$ that is isometrically isomorphic to the corresponding X_j.

PROOF. The map $x \mapsto (x, 0, \ldots, 0)$ is clearly an isometric isomorphism from X_1 onto X_1'. If $\big((x^{(k)}, 0, \ldots, 0)\big)$ is a sequence in X_1' and $(x_1, \ldots, x_n)$ is a member of $X_1 \oplus \cdots \oplus X_n$ such that $\|(x^{(k)}, 0, \ldots, 0) - (x_1, \ldots, x_n)\| \to 0$, then it is easy to see that $x_2, \ldots, x_n$ are all zero, from which it follows that X_1' is closed in $X_1 \oplus \cdots \oplus X_n$. Analogous arguments work for the other X_j's. ∎

The following two propositions give important "algebraic" properties of direct sums. The first is a generalized commutative and associative law, while the second serves as a cancellation law.

1.8.4 Proposition. *Let $X_1, \ldots, X_n$ be normed spaces. If two direct sums are each formed by permuting and associating the terms of $X_1 \oplus \cdots \oplus X_n$, then those two direct sums are isometrically isomorphic.*

PROOF. It is enough to prove that a direct sum formed by permuting and associating the terms of $X_1 \oplus \cdots \oplus X_n$ is isometrically isomorphic to $X_1 \oplus \cdots \oplus X_n$. An example will tell most of the story. Suppose for the moment that $n = 5$, and consider the direct sum $\bigl(X_4 \oplus (X_5 \oplus X_2)\bigr) \oplus X_3 \oplus X_1$ formed from $X_1 \oplus X_2 \oplus X_3 \oplus X_4 \oplus X_5$ by permuting and associating its terms. It is easy to check that the map

$$\Bigl(\bigl(x_4, (x_5, x_2)\bigr), x_3, x_1\Bigr) \mapsto (x_1, x_2, x_3, x_4, x_5)$$

from $\bigl(X_4 \oplus (X_5 \oplus X_2)\bigr) \oplus X_3 \oplus X_1$ onto $X_1 \oplus X_2 \oplus X_3 \oplus X_4 \oplus X_5$ is an isometric isomorphism. For any n and any direct sum Y formed by permuting and associating the terms of $X_1 \oplus \cdots \oplus X_n$, the obvious analog of the above map is clearly an isometric isomorphism from Y onto $X_1 \oplus \cdots \oplus X_n$. ■

1.8.5 Proposition. *Let $X_1, \ldots, X_n$ be normed spaces. Let $\{1, \ldots, n\}$ be partitioned into two nonempty sets $\{j_1, \ldots, j_p\}$ and $\{k_1, \ldots, k_q\}$ and let*

$$X_{j_1,\ldots,j_p} = \{\,(x_1, \ldots, x_n) : (x_1, \ldots, x_n) \in X_1 \oplus \cdots \oplus X_n,\quad x_{k_1}, \ldots, x_{k_q} \text{ are all zero}\}.$$

Then $(X_1 \oplus \cdots \oplus X_n)/X_{j_1,\ldots,j_p}$ is isometrically isomorphic to $X_{k_1} \oplus \cdots \oplus X_{k_q}$.

PROOF. It is clear that $X_{j_1,\ldots,j_p}$ is a subspace of $X_1 \oplus \cdots \oplus X_n$. If an element $(z_1, \ldots, z_n)$ of $X_1 \oplus \cdots \oplus X_n$ is the limit of a sequence of elements of $X_{j_1,\ldots,j_p}$, then $z_{k_1}, \ldots, z_{k_q}$ must all be zero, so the subspace $X_{j_1,\ldots,j_p}$ is closed. Thus, the quotient space $(X_1 \oplus \cdots \oplus X_n)/X_{j_1,\ldots,j_p}$ can be formed.

Define $T\colon (X_1 \oplus \cdots \oplus X_n)/X_{j_1,\ldots,j_p} \to X_{k_1} \oplus \cdots \oplus X_{k_q}$ by the formula $T\bigl((x_1, \ldots, x_n) + X_{j_1,\ldots,j_p}\bigr) = (x_{k_1}, \ldots, x_{k_q})$. Since $(x_1, \ldots, x_n) + X_{j_1,\ldots,j_p} = (y_1, \ldots, y_n) + X_{j_1,\ldots,j_p}$ if and only if $x_{k_m} = y_{k_m}$ when $m = 1, \ldots, q$, the map T is well-defined. It is easy to check that T is linear, one-to-one, and onto $X_{k_1} \oplus \cdots \oplus X_{k_q}$. If $(x_1, \ldots, x_n) + X_{j_1,\ldots,j_p} \in (X_1 \oplus \cdots \oplus X_n)/X_{j_1,\ldots,j_p}$, then

$$\begin{aligned}
&\|(x_1, \ldots, x_n) + X_{j_1,\ldots,j_p}\| \\
&\quad = d\bigl((0, \ldots, 0), (x_1, \ldots, x_n) + X_{j_1,\ldots,j_p}\bigr) \\
&\quad = \inf\Bigl\{\Bigl(\sum_{m=1}^{n} \|y_m\|^2\Bigr)^{1/2} : (y_1, \ldots, y_n) \in X_1 \oplus \cdots \oplus X_n, \\
&\qquad\qquad y_{k_1} = x_{k_1}, \ldots, y_{k_q} = x_{k_q}\Bigr\} \\
&\quad = \Bigl(\sum_{m=1}^{q} \|x_{k_m}\|^2\Bigr)^{1/2} \\
&\quad = \|(x_{k_1}, \ldots, x_{k_q})\| \\
&\quad = \bigl\|T\bigl((x_1, \ldots, x_n) + X_{j_1,\ldots,j_p}\bigr)\bigr\|,
\end{aligned}$$

so T is the desired isometric isomorphism. ■

With all notation as in the statement of the preceding proposition, it is easy to see that $X_{j_1,\dots,j_p}$ is isometrically isomorphic to $X_{j_1} \oplus \cdots \oplus X_{j_p}$. If spaces that are isometrically isomorphic are considered to be the same, then the proposition says that

$$\frac{X_1 \oplus \cdots \oplus X_n}{X_{j_1} \oplus \cdots \oplus X_{j_p}} = X_{k_1} \oplus \cdots \oplus X_{k_q},$$

which is why the proposition can be called a cancellation law. This cancellation law would be more appealing visually if the symbol $\otimes$ were used to separate the terms of a direct sum, but that symbol is reserved for tensor products. Because of their product-like behavior and their relationship to Cartesian products, direct sums are sometimes called direct products, but then so are tensor products. For this reason, the term *direct product* will not be used in this book.

A metric space is called *topologically complete* if its metric is equivalent to a complete metric, that is, if the topology of the space is induced by *some* complete metric, even if the *given* metric is not complete. If $X_1, \dots, X_n$ are metric spaces, then the product topology of $X_1 \times \cdots \times X_n$ is topologically complete if and only if each X_j is topologically complete; see, for example, [65]. This suggests the following theorem, though it does not prove it. As was shown in Exercise 1.42, it is quite possible for an incomplete metric space to be topologically complete.[5]

1.8.6 Theorem. *Let $X_1, \dots, X_n$ be normed spaces. Then $X_1 \oplus \cdots \oplus X_n$ is a Banach space if and only if each X_j is a Banach space.*

PROOF. Suppose that each X_j is a Banach space. Let $\big((x_1^{(k)}, \dots, x_n^{(k)})\big)_{k=1}^{\infty}$ be a Cauchy sequence in $X_1 \oplus \cdots \oplus X_n$. Then the formula for the direct sum norm assures that each of the sequences $(x_j^{(k)})$ such that $j = 1, \dots, n$ is also Cauchy, and so there is an element $(x_1, \dots, x_n)$ of $X_1 \oplus \cdots \oplus X_n$ such that $\lim_k x_j^{(k)} = x_j$ when $j = 1, \dots, n$. Thus, the sequence $\big((x_1^{(k)}, \dots, x_n^{(k)})\big)$ converges to $(x_1, \dots, x_n)$, and so $X_1 \oplus \cdots \oplus X_n$ is a Banach space.

Conversely, suppose that $X_1 \oplus \cdots \oplus X_n$ is a Banach space. By Proposition 1.8.3, there are closed, hence complete, subspaces $X_1', \dots, X_n'$ of $X_1 \oplus \cdots \oplus X_n$ that are isometrically isomorphic to $X_1, \dots, X_n$ respectively, which assures that each X_j is a Banach space. ■

Some notation is needed before proceeding. Suppose that V is a vector space and that $A_1, \dots, A_n$ are subsets of V. Then the *(algebraic) sum* of $A_1, \dots, A_n$ is $\{\, a_1 + \cdots + a_n : a_j \in A_j \text{ for each } j \,\}$ and is denoted by

[5] Actually, it is not possible for an incomplete *normed* space to be topologically complete, as will be shown in Corollary 2.3.19. However, that fact, which first appeared in a 1952 paper by Victor Klee [135], is much deeper than what is needed to prove Theorem 1.8.6.

$A_1 + \cdots + A_n$ or $\sum_{j=1}^n A_j$. Notice that this is just the natural extension of the notion of the sum of two subsets of a vector space. It is easy to see that if $A_1, \ldots, A_n$ are subspaces of V, then so is their sum. By convention, the sum of an empty collection of subspaces of V is $\{0\}$.

Let X and Y be normed spaces and let the subsets $\{(x, 0) : x \in X\}$ and $\{(0, y) : y \in Y\}$ of $X \oplus Y$ be denoted by X' and Y' respectively. By Proposition 1.8.3, the sets X' and Y' are closed subspaces of $X \oplus Y$ isometrically isomorphic to X and Y respectively. Notice that $X' + Y' = X \oplus Y$ and that $X' \cap Y' = \{(0,0)\}$. In vector space terminology, the vector space underlying $X \oplus Y$ would be called the *algebraic internal direct sum* of its subspaces X' and Y'. More generally, a vector space X is said to be the algebraic internal direct sum of its subspaces $M_1, \ldots, M_n$ if $\sum_k M_k = X$ and $M_j \cap \sum_{k \neq j} M_k = \{0\}$ when $j = 1, \ldots, n$. The following is a useful characterization of these objects.

1.8.7 Proposition. *Suppose that X is a vector space and that $M_1, \ldots, M_n$ are subspaces of X. Then the following are equivalent.*

(a) *The space X is the algebraic internal direct sum of $M_1, \ldots, M_n$.*

(b) *For every x in X, there are unique elements $m_1(x), \ldots, m_n(x)$ of $M_1, \ldots, M_n$ respectively such that $x = \sum_k m_k(x)$.*

PROOF. Suppose first that (a) holds. Since $X = \sum_k M_k$, there must be elements $m_1, \ldots, m_n$ of X such that $m_k \in M_k$ for each k and $x = \sum_k m_k$. Suppose that $m_1', \ldots, m_n' \in X$, that $m_k' \in M_k$ for each k, and that $x = \sum_k m_k'$. If $1 \leq j \leq n$, then

$$m_j - m_j' = \sum_{k \neq j} (m_k' - m_k) \in M_j \cap \sum_{k \neq j} M_k = \{0\},$$

so $m_j = m_j'$. This shows that (a) $\Rightarrow$ (b).

Suppose conversely that (b) holds. It is immediate that $\sum_k M_k = X$. If $1 \leq j \leq n$ and $x \in M_j \cap \sum_{k \neq j} M_k$, then there are members $m_1, \ldots, m_n$ of X such that $m_k \in M_k$ for each k and $x = m_j = \sum_{k \neq j} m_k$, which implies that $m_j + \sum_{k \neq j}(-m_k) = 0 = 0 + \sum_{k \neq j} 0$ and therefore that $m_1 = \cdots = m_n = 0$, which in turn implies that $x = 0$. Therefore $M_j \cap \sum_{k \neq j} M_k = \{0\}$ when $j = 1, \ldots, n$, so X is the algebraic internal direct sum of $M_1, \ldots, M_n$. This proves that (b) $\Rightarrow$ (a). ∎

The internal direct sums that are of the most importance in the theory of normed spaces have an additional restriction on the subspaces used to form them.

1.8.8 Definition. Suppose that $M_1, \ldots, M_n$ are closed subspaces of a normed space X such that $\sum_k M_k = X$ and $M_j \cap \sum_{k \neq j} M_k = \{0\}$ when $j = 1, \ldots, n$. Then the normed space X is the *(internal) direct sum* of $M_1, \ldots, M_n$.

That is, a normed space X is the internal direct sum of its subspaces $M_1, \ldots, M_n$ if and only if each of these subspaces is closed and X is their algebraic internal direct sum.

1.8.9 Example. In $C[0,1]$, let M_1 be the collection of all constant functions, let M_2 be the collection of all linear functions f such that $f(0) = 0$, and let M_3 be the collection of all f such that $f(0) = f(1) = 0$. It is easy to check that M_1, M_2, and M_3 are closed subspaces of $C[0,1]$ and that $C[0,1]$ is the internal direct sum of the three.

Notice that the use of the words *external* and *internal* is optional when referring to normed space direct sums. This means that a statement such as "the normed space X is the direct sum of Y and Z" could be ambiguous. Normally, the context does make it clear which type of direct sum is being considered. Even when it does not, the two types of direct sums are so closely related that problems usually do not arise. This relationship is the subject of the next proposition.

1.8.10 Proposition.

(a) *If $X_1, \ldots, X_n$ are normed spaces and $X = X_1 \oplus \cdots \oplus X_n$, then X has closed subspaces $X_1', \ldots, X_n'$ such that X is the internal direct sum of $X_1', \ldots, X_n'$ and each X_j is isometrically isomorphic to the corresponding X_j'.*

(b) *If X is a Banach space that is the internal direct sum of its closed subspaces $M_1, \ldots, M_n$, then $X \cong M_1 \oplus \cdots \oplus M_n$.*

PROOF. For (a), let $X_1', \ldots, X_n'$ be as in Proposition 1.8.3. It is easy to check that these closed subspaces of X do what is needed. For (b), suppose that X is a Banach space that is the internal direct sum of its closed subspaces $M_1, \ldots, M_n$. Since each M_j is a Banach space, so is $M_1 \oplus \cdots \oplus M_n$. The map $T\colon M_1 \oplus \cdots \oplus M_n \to X$ given by the formula $T(m_1, \ldots, m_n) = m_1 + \cdots + m_n$ is clearly linear and onto X. By Corollary 1.4.18, the ℓ_1^n and ℓ_2^n norms on $\mathbb{F}^n$ are equivalent, from which it follows that there is a positive constant c such that

$$\|T(m_1, \ldots, m_n)\| \le \sum_{j=1}^{n} \|m_j\| \le c\Big(\sum_{j=1}^{n} \|m_j\|^2\Big)^{1/2} = c\,\|(m_1, \ldots, m_n)\|$$

for each member $(m_1, \ldots, m_n)$ of $M_1 \oplus \cdots \oplus M_n$. Thus, the operator T is bounded. If $T(m_1, \ldots, m_n) = 0$, then $m_1 = \cdots = m_n = 0$ by Proposition 1.8.7. It follows that T is one-to-one. By Corollary 1.6.6, the operator T is an isomorphism from $M_1 \oplus \cdots \oplus M_n$ onto X. ■

Part (b) of the preceding proposition is *not* in general true for incomplete normed spaces. See Exercise 1.95.

The notion of the direct sum of normed spaces leads in a very natural way to the notion of the direct sum of linear operators between normed spaces.

1.8.11 Definition. Suppose that $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ are normed spaces and that T_j is a linear operator from X_j into Y_j when $j = 1, \ldots, n$. Then the *direct sum* of $T_1, \ldots, T_n$ is the map

$$T_1 \oplus \cdots \oplus T_n \colon\ X_1 \oplus \cdots \oplus X_n \to Y_1 \oplus \cdots \oplus Y_n$$

defined by letting

$$T_1 \oplus \cdots \oplus T_n\ (x_1, \ldots, x_n) = (T_1 x_1, \ldots, T_n x_n)$$

whenever $(x_1, \ldots, x_n) \in X_1 \oplus \cdots \oplus X_n$.

It is easy to check that the map $T_1 \oplus \cdots \oplus T_n$ of the preceding definition is linear. As the next theorem indicates, special properties possessed by each T_j tend to be inherited by $T_1 \oplus \cdots \oplus T_n$.

1.8.12 Theorem. *Suppose that $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ are normed spaces and that T_j is a linear operator from X_j into Y_j when $j = 1, \ldots, n$. Then $T_1 \oplus \cdots \oplus T_n$ is bounded if and only if each T_j is bounded. If $T_1 \oplus \cdots \oplus T_n$ is bounded, then $\|T_1 \oplus \cdots \oplus T_n\| = \max\{\|T_1\|, \ldots, \|T_n\|\}$. Furthermore, the operator $T_1 \oplus \cdots \oplus T_n$ is one-to-one, or onto, or an isomorphism, or an isometric isomorphism, if and only if $T_1, \ldots, T_n$ all have that same property.*

PROOF. Suppose that T_1 is unbounded. Then

$$\begin{aligned}
\sup\{\, \|T_1 \oplus \cdots \oplus T_n(x_1, \ldots, x_n)\| : (x_1, \ldots, x_n) \in B_{X_1 \oplus \cdots \oplus X_n} \,\}& \\
\geq \sup\{\, \|T_1 \oplus \cdots \oplus T_n(x_1, 0, \ldots, 0)\| : x_1 \in B_{X_1} \,\}& \\
= \sup\{\, \|T_1 x_1\| : x_1 \in B_{X_1} \,\}& \\
= +\infty,&
\end{aligned}$$

so $T_1 \oplus \cdots \oplus T_n$ is unbounded. Similarly, the operator $T_1 \oplus \cdots \oplus T_n$ is unbounded whenever any T_j is unbounded. Conversely, suppose that each T_j is bounded. For each member $(x_1, \ldots, x_n)$ of $X_1 \oplus \cdots \oplus X_n$,

$$\begin{aligned}
\|T_1 \oplus \cdots \oplus T_n(x_1, \ldots, x_n)\| &= \Big(\sum_{j=1}^n \|T_j x_j\|^2 \Big)^{1/2} \\
&\leq \Big(\sum_{j=1}^n (\|T_j\|\, \|x_j\|)^2 \Big)^{1/2} \\
&\leq \max\{\|T_1\|, \ldots, \|T_n\|\} \Big(\sum_{j=1}^n \|x_j\|^2 \Big)^{1/2} \\
&= \max\{\|T_1\|, \ldots, \|T_n\|\} \|(x_1, \ldots, x_n)\|,
\end{aligned}$$

which implies that $T_1 \oplus \cdots \oplus T_n$ is bounded and that $\|T_1 \oplus \cdots \oplus T_n\| \le \max\{\|T_1\|, \ldots, \|T_n\|\}$. Notice next that whenever $x_1 \in B_{X_1}$,

$$\|T_1 \oplus \cdots \oplus T_n\| \ge \|T_1 \oplus \cdots \oplus T_n(x_1, 0, \ldots, 0)\| = \|T_1 x_1\|,$$

so $\|T_1 \oplus \cdots \oplus T_n\| \ge \|T_1\|$. It follows similarly that $\|T_1 \oplus \cdots \oplus T_n\| \ge \|T_j\|$ when $j = 2, \ldots, n$, so $\|T_1 \oplus \cdots \oplus T_n\| \ge \max\{\|T_1\|, \ldots, \|T_n\|\}$. Therefore $\|T_1 \oplus \cdots \oplus T_n\| = \max\{\|T_1\|, \ldots, \|T_n\|\}$.

It is easy to see that $T_1 \oplus \cdots \oplus T_n$ is one-to-one, or onto, or an isometric isomorphism, if and only if each T_j has the corresponding property. All that remains to be proved is that $T_1 \oplus \cdots \oplus T_n$ is an isomorphism if and only if each T_j is so. The failure of one of the T_j's to be one-to-one would prohibit both $T_1 \oplus \cdots \oplus T_n$ and that T_j from being isomorphisms, so it can be assumed that each T_j is one-to-one. Since $T_1 \oplus \cdots \oplus T_n$ can be viewed as an operator onto the normed space $T(X_1) \oplus \cdots \oplus T(X_n)$, no harm comes from also assuming that each T_j is onto the corresponding Y_j. It is then easy to see that $(T_1 \oplus \cdots \oplus T_n)^{-1} = T_1^{-1} \oplus \cdots \oplus T_n^{-1}$, and so $(T_1 \oplus \cdots \oplus T_n)^{-1}$ is bounded if and only if each T_j^{-1} is bounded. Thus, the operator $T_1 \oplus \cdots \oplus T_n$ is an isomorphism exactly when each T_j is an isomorphism. ■

1.8.13 Corollary. *Suppose that $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ are normed spaces. If $X_j \cong Y_j$ when $j = 1, \ldots, n$, then $X_1 \oplus \cdots \oplus X_n \cong Y_1 \oplus \cdots \oplus Y_n$. If X_j is isometrically isomorphic to Y_j when $j = 1, \ldots, n$, then $X_1 \oplus \cdots \oplus X_n$ is isometrically isomorphic to $Y_1 \oplus \cdots \oplus Y_n$.*

As an application of the results of this section, it will now be shown that a bounded linear operator from a subspace of a Banach space X into a normed space Y can always be extended to a bounded linear operator from X into Y if the subspace has the following property.

1.8.14 Definition. A subspace M of a normed space X is *complemented* in X if it is closed in X and there is a closed subspace N of X such that X is the internal direct sum of M and N, in which case the subspace N is said to be *complementary* to M.

The study of complemented subspaces of Banach spaces will be taken up in earnest in Section 3.2, where it will be shown that deciding whether or not a closed subspace M of a Banach space X is complemented is equivalent to deciding whether or not there is a bounded linear operator P mapping X onto M such that $P(Px) = Px$ for every x in X. An operator with the properties just described is called a *bounded projection* from X onto M. The following argument is based on the construction of such a projection.

Suppose that T is a bounded linear operator from a complemented subspace M of a Banach space X into a normed space Y. Let N be a subspace complementary to M and let $S\colon M \oplus N \to X$ be defined by the formula $S(m,n) = m + n$. By the proof of Proposition 1.8.10 (b), the map S is an isomorphism from $M \oplus N$ onto X, so S^{-1} is an isomorphism from X onto $M \oplus N$. Notice that $S^{-1}m = (m,0)$ for each m in M. Let $R(m,n) = m$ for each member (m,n) of $M \oplus N$. Since R is a bounded linear operator from $M \oplus N$ onto M, the product RS^{-1} is a bounded linear operator from X onto M; call it P. Notice that $Pm = m$ for each m in M (and therefore $P(Px) = Px$ for each x in X). It follows immediately that TP is a bounded linear operator from X into Y that agrees with T on M.

More can be said if T is a bounded linear functional on a subspace M of a normed space X. In that case, it turns out that T can always be extended to a bounded linear functional on all of X, even if X is not a Banach space and M is not complemented or, for that matter, even closed in X. If T_e is such an extension, then

$$\|T_e\| = \sup\{\, \|T_e x\| : x \in B_X \,\} \geq \sup\{\, \|Tm\| : m \in B_M \,\} = \|T\|,$$

but it turns out that T_e can even be selected so that $\|T_e\| = \|T\|$. That is the content of the main result of the next section.

Exercises

1.88 Suppose that $X_1, \dots, X_n$ are normed spaces and that $1 \leq p \leq \infty$. To avoid an impending notational conflict, let the norm of ℓ_p^n be denoted by $\|\cdot\|_{\ell_p^n}$ instead of $\|\cdot\|_p$. Show that the formula

$$\|(x_1, \dots, x_n)\|_p = \big\|\big(\|x_1\|, \dots, \|x_n\|\big)\big\|_{\ell_p^n}$$

defines a norm on the vector space sum of $X_1, \dots, X_n$ equivalent to the usual direct sum norm. Notice that this formula produces the usual direct sum norm when $p = 2$.

1.89 Prove or disprove the following generalization of the preceding exercise: If $/\cdot/$ is a norm on $\mathbb{F}^n$, then the formula

$$//(x_1, \dots, x_n)// = /\big(\|x_1\|, \dots, \|x_n\|\big)/$$

must define a norm on the vector space sum of $X_1, \dots, X_n$ equivalent to the usual direct sum norm.

1.90 Suppose that X is c_0 or ℓ_p, where $1 \leq p \leq \infty$. Prove that $X \oplus \mathbb{F} \cong X$. Conclude that $X \oplus Y \cong X$ whenever Y is a finite-dimensional normed space.

1.91 Suppose that X is c_0 or ℓ_p, where $1 \leq p \leq \infty$. Prove that $X \oplus X \cong X$.

1.92 Prove that $L_p[0,1] \oplus L_p[0,1] \cong L_p[0,1]$ when $1 \leq p \leq \infty$. One way to do this requires some knowledge of how to change variables in the Lebesgue integral. See, for example, [202, pp. 153–155].

1.93 (a) Suppose that X, Y, and Z are normed spaces such that $Y \oplus Y \cong Y$ and Z is the direct sum of X with a finite positive number of copies of Y. Prove that $Z \cong X \oplus Y$.

(b) Suppose that X and Y are Banach spaces such that $Y \oplus Y \cong Y$ and X has a complemented subspace isomorphic to Y. Prove that $X \oplus Y \cong X$, and therefore that every direct sum of X with a finite number of copies of Y is isomorphic to X.

1.94 (a) Prove that $\ell_2 \oplus \ell_2$ is isometrically isomorphic to ℓ_2.

(b) Suppose that X is a normed space and that Y is the direct sum of X with a finite positive number of copies of ℓ_2. Prove that Y is isometrically isomorphic to $X \oplus \ell_2$.

1.95 (a) Conclude from Exercise 1.84 that Proposition 1.8.10 (b) does not always hold if "Banach space" is replaced by "normed space."

(b) Let X_1 and X_2 be normed spaces such that X_j is the internal direct sum of its closed subspaces M_j and N_j when $j = 1, 2$. Suppose that M_1 and M_2 are isometrically isomorphic, as are N_1 and N_2. Conclude from (a) that X_1 and X_2 need not even be isomorphic.

(c) In part (b), suppose that X_1 and X_2 are Banach spaces. Can you then conclude that they are isomorphic? How about isometrically isomorphic?

1.96 Suppose that the direct sum norm were defined to be any of the equivalent norms of Exercise 1.88. Prove that if $T_1, \ldots, T_n$ are bounded linear operators, then the norm of their direct sum is still given by the formula $\|T_1 \oplus \cdots \oplus T_n\| = \max\{\|T_1\|, \ldots, \|T_n\|\}$.

1.97 In this exercise, no result from Section 1.4 about finite-dimensional vector spaces may be used. (However, feel free to use Proposition 1.8.10. The application of Corollary 1.4.18 in the proof of that proposition can be replaced by an application of Cauchy's inequality.) Let n be a positive integer and let X be a normed space of dimension n. Prove that X is the internal direct sum of n subspaces each isometrically isomorphic to $\mathbb{F}$. Use this to prove that all n-dimensional Banach spaces over $\mathbb{F}$ are isomorphic to each other. (Notice that this is a weakened form of Corollary 1.4.16.)

1.98 Suppose that X and Y are Banach spaces and that there is a bounded linear operator from X onto Y whose kernel is complemented in X. Prove that X has a complemented subspace isomorphic to Y.

1.99 Let T be a bounded linear operator from a Banach space X into a finite-dimensional normed space Y. The purpose of this exercise is to show that T is, in a sense, the direct sum of a Banach space isomorphism and a zero operator.

(a) Define a subspace M of X as follows. First show that $X/\ker(T)$ and $T(X)$ have the same finite dimension n. If $T(X) = \{0\}$, then let $M = \{0\}$. If $T(X) \neq \{0\}$, then let $x_1 + \ker(T), \ldots, x_n + \ker(T)$ be a basis for $X/\ker(T)$ and let $M = \langle x_1, \ldots, x_n \rangle$. Show that in the second case the vectors $x_1, \ldots, x_n$ are linearly independent, and that in either case M is a closed subspace of X such that X is the internal direct sum of M and $\ker(T)$.

(b) Let $\mathfrak{B}_1$ be a basis for $T(X)$. Show that there is a collection $\mathfrak{B}_2$ of linearly independent vectors in Y such that $\langle\mathfrak{B}_1\rangle \cap \langle\mathfrak{B}_2\rangle = \{0\}$ and $\mathfrak{B}_1 \cup \mathfrak{B}_2$ is a basis for Y. Let $Z = \langle\mathfrak{B}_2\rangle$. Show that Y is the internal direct sum of $T(X)$ and Z.

(c) Define $T_1 \colon M \to T(X)$ by the formula $T_1(m) = T(m)$. Show that T_1 is an isomorphism from M onto $T(X)$.

(d) Define $I_1 \colon X \to M \oplus \ker(T)$ and $I_2 \colon T(X) \oplus Z \to Y$ by letting $I_1(m+k) = (m,k)$ whenever $m \in M$ and $k \in \ker(T)$, and $I_2(w,z) = w + z$ whenever $(w,z) \in T(X) \oplus Z$. Conclude from the proof of Proposition 1.8.10 (b) that I_1 and I_2 are isomorphisms from X onto $M \oplus \ker(T)$ and from $T(X) \oplus Z$ onto Y respectively.

(e) Let T_2 be the zero operator from $\ker(T)$ into Z. Show that $T = I_2(T_1 \oplus T_2)I_1$. (If I_1 and I_2 are treated as ways of identifying X with $M \oplus \ker(T)$ and $T(X) \oplus Z$ with Y, then T "is" $T_1 \oplus T_2$.)

1.9 The Hahn-Banach Extension Theorems

Suppose that X and Y are normed spaces and that M is a subspace of X. It is often important to know whether a bounded linear operator from M into Y can be extended to a bounded linear operator from all of X into Y. When this is possible, it can also be important to know whether it can be done without increasing the norm of the operator. There is no problem doing so when Y is a Banach space and M is dense in X.

1.9.1 Theorem. *Suppose that M is a dense subspace of a normed space X, that Y is a Banach space, and that $T_0 \colon M \to Y$ is a bounded linear operator. Then there is a unique continuous function $T \colon X \to Y$ that agrees with T_0 on M. This function T is a bounded linear operator, and $\|T\| = \|T_0\|$. If T_0 is an isomorphism or isometric isomorphism, then T has that same property.*

PROOF. If (x_n) is a Cauchy sequence in M, then (T_0x_n) is a Cauchy sequence in Y since $\|T_0x_{n_1} - T_0x_{n_2}\| \le \|T_0\|\,\|x_{n_1} - x_{n_2}\|$ whenever $n_1, n_2 \in \mathbb{N}$, so (T_0x_n) converges. Also, if (w_n) and (x_n) are sequences in M that converge to the same limit in X, then $\|T_0w_n - T_0x_n\| \le \|T_0\|\,\|w_n - x_n\| \to 0$ as $n \to \infty$, so $\lim_n T_0w_n = \lim_n T_0x_n$. Thus, a function T can be unambiguously defined on X by the formula $T(x) = \lim_n T_0x_n$, where (x_n) is any sequence in M that converges to x.

For each m in M, the constant sequence (m_n) with m as each term converges to m, implying that $T(m) = \lim_n T_0m_n = T_0m$, so T and T_0 agree on M. If $w, x \in X$, $\alpha \in \mathbb{F}$, and (w_n) and (x_n) are sequences in M

converging to w and x respectively, then $\alpha w_n + x_n \to \alpha w + x$, so

$$\begin{aligned} T(\alpha w + x) &= \lim_n T_0(\alpha w_n + x_n) \\ &= \alpha \lim_n T_0 w_n + \lim_n T_0 x_n \\ &= \alpha T(w) + T(x). \end{aligned}$$

It follows that T is linear. Since B_M is dense in B_X, it is clear from the definition of T that $\sup\{ \|Tx\| : x \in B_X \} = \sup\{ \|T_0 m\| : m \in B_M \}$, so T is bounded and has the same norm as T_0. The uniqueness assertion follows from the fact that two continuous functions from one metric space into another that agree on a dense subset of their domain must agree on the whole domain.

Finally, suppose that T_0 is an isomorphism. By Proposition 1.4.14 (a), there are positive constants s and t such that $s\|m\| \le \|T_0 m\| \le t\|m\|$ whenever $m \in M$. Notice that T_0 is an isometric isomorphism if and only if this holds when $s = t = 1$. It is an easy consequence of the density of M in X and the continuity of T that $s\|x\| \le \|Tx\| \le t\|x\|$ whenever $x \in X$. It follows that T is an isomorphism, and is an isometric isomorphism if T_0 is an isometric isomorphism. ∎

The requirement that Y be complete cannot be omitted, as is shown in Exercise 1.100. Problems also arise if the requirement that M be dense in X is removed. For example, it turns out that when c_0 is treated in the usual way as a subspace of ℓ_∞, the identity operator on c_0 cannot be extended to a bounded linear operator from ℓ_∞ onto c_0. See Corollary 3.2.21.

The main purpose of this section is to show that a bounded linear functional on a subspace of a normed space can always be extended to a bounded linear functional on the entire space without increasing its norm. The plan is to prove this for real normed spaces, then pass from the real case to the complex. To accomplish this, some preliminary facts about real-linear functionals on complex vector spaces are needed.

1.9.2 Definition. Let X be a complex vector space. A *real-linear functional* on X is a real-valued function f on X such that if $x, y \in X$ and $\alpha \in \mathbb{R}$, then $f(x + y) = f(x) + f(y)$ and $f(\alpha x) = \alpha f(x)$.

Of course, every complex vector space X is also a real vector space X_r when multiplication of vectors by scalars is restricted to $\mathbb{R} \times X$. A real-linear functional on X is just a linear functional in the usual sense on X_r. For the moment, a linear functional in the usual sense on X will be called a *complex-linear functional* on X.

As usual, the real and imaginary parts of a complex number α will be denoted by $\operatorname{Re}\alpha$ and $\operatorname{Im}\alpha$ respectively.

1.9.3 Proposition. *Let X be a complex vector space and let X_r be the corresponding real vector space.*

(a) *If f is a complex-linear functional on X and u is the real part of f, then u is a real-linear functional on X, and $f(x) = u(x) - iu(ix)$ whenever $x \in X$.*

(b) *If u is a real-linear functional on X, then there is a unique complex-linear functional f on X such that u is the real part of f. This functional f is given by the formula from* (a).

(c) *Suppose that X is a complex normed space and that f is a complex-linear functional on X with real part u. Then f is a bounded linear functional on X if and only if u is a bounded linear functional on X_r. If f is bounded, then $\|f\| = \|u\|$.*

PROOF. Let f be a complex-linear functional on X and let u be its real part. Since $\alpha = \operatorname{Re}(\alpha) - i\operatorname{Re}(i\alpha)$ for every complex number α, it follows that

$$f(x) = \operatorname{Re}\big(f(x)\big) - i\operatorname{Re}\big(if(x)\big) = \operatorname{Re}\big(f(x)\big) - i\operatorname{Re}\big(f(ix)\big) = u(x) - iu(ix)$$

whenever $x \in X$. If $x, y \in X$ and α is a real scalar, Then

$$\begin{aligned} u(\alpha x + y) &= \operatorname{Re}\big(\alpha f(x) + f(y)\big) \\ &= \operatorname{Re}\big(\alpha u(x) - i\alpha u(ix) + u(y) - iu(iy)\big) \\ &= \alpha u(x) + u(y), \end{aligned}$$

so u is a real-linear functional on X. This proves (a).

For (b), let u be a real-linear functional on X. It follows from (a) that the function f defined by the formula $f(x) = u(x) - iu(ix)$ is the only possible candidate for a complex-linear functional on X with real part u. If $x, y \in X$ and α is a complex scalar, then

$$\begin{aligned} f(\alpha x + y) &= u(\alpha x + y) - iu(i\alpha x + iy) \\ &= u\big(\operatorname{Re}(\alpha)x\big) + u\big(i\operatorname{Im}(\alpha)x\big) + u(y) \\ &\qquad - iu\big(i\operatorname{Re}(\alpha)x\big) - iu\big(-\operatorname{Im}(\alpha)x\big) - iu(iy) \\ &= \operatorname{Re}(\alpha)u(x) + \operatorname{Im}(\alpha)u(ix) + u(y) \\ &\qquad - i\operatorname{Re}(\alpha)u(ix) + i\operatorname{Im}(\alpha)u(x) - iu(iy) \\ &= \alpha u(x) - i\alpha u(ix) + u(y) - iu(iy) \\ &= \alpha f(x) + f(y), \end{aligned}$$

which shows that f is a complex-linear functional on X. This finishes the proof of (b).

Finally, suppose that X is a complex normed space and that f is a complex-linear functional on X with real part u. For each x in X, there is a scalar α_x such that $|\alpha_x| = 1$ and $\alpha_x f(x)$ is a nonnegative real number. Notice that $|u(x)| \le |f(x)| = f(\alpha_x x) = u(\alpha_x x)$ whenever $x \in X$. Therefore

$$\begin{aligned}\sup\{\,|u(x)| : x \in B_{X_r}\,\} &\le \sup\{\,|f(x)| : x \in B_X\,\}\\ &= \sup\{\,u(\alpha_x x) : x \in B_{X_r}\,\}\\ &\le \sup\{\,|u(x)| : x \in B_{X_r}\,\},\end{aligned}$$

so $\sup\{\,|u(x)| : x \in B_{X_r}\,\} = \sup\{\,|f(x)| : x \in B_X\,\}$. Part (c) follows immediately from this. ■

It is time for the main results of this section. There is a body of related facts, collectively called the *Hahn-Banach theorem,* that includes Theorems 1.9.5 and 1.9.6 and Proposition 1.9.15 from this section as well as Theorems 2.2.19, 2.2.26, and 2.2.28 in the next chapter. The common theme of all of them is that under certain conditions a vector space always has a large enough supply of well-behaved linear functionals to accomplish certain tasks.

1.9.4 Definition. Let p be a real-valued function on a vector space X. Then p is *positive-homogeneous* if $p(tx) = tp(x)$ whenever $t > 0$ and $x \in X$, and is *(finitely) subadditive* if $p(x+y) \le p(x) + p(y)$ whenever $x, y \in X$. If p has both properties, then it is said to be a *sublinear functional.*

Notice that every seminorm, and in particular every norm, is a sublinear functional, as is every linear functional on a real vector space.

1.9.5 The Vector Space Version of the Hahn-Banach Extension Theorem. (S. Banach, 1929 [11]). *Suppose that p is a sublinear functional on a real vector space X and that f_0 is a linear functional on a subspace Y of X such that $f_0(y) \le p(y)$ whenever $y \in Y$. Then there is a linear functional f on all of X such that the restriction of f to Y is f_0 and $f(x) \le p(x)$ whenever $x \in X$. That is, the linear functional f_0 can be extended to a linear functional on X that is still dominated by p.*

PROOF. The first step is to show that if Y is not all of X, then there is a linear extension f_1 of f_0 to a subspace of X larger than Y such that f_1 is still dominated by p. Suppose that $x_1 \in X \setminus Y$. Let $Y_1 = Y + \langle\{x_1\}\rangle$, a subspace of X that properly includes Y. Notice that if $y + tx_1 = y' + t'x_1$ where $y, y' \in Y$ and $t, t' \in \mathbb{R}$, then $(t - t')x_1 = y' - y \in Y$, and so $t = t'$ and $y = y'$. Thus, each member of Y_1 can be expressed in the form $y + tx_1$, where $y \in Y$ and $t \in \mathbb{R}$, in exactly one way. Whenever $y_1, y_2 \in Y$,

$$\begin{aligned}f_0(y_1) + f_0(y_2) &= f_0(y_1 + y_2)\\ &\le p(y_1 - x_1 + y_2 + x_1)\\ &\le p(y_1 - x_1) + p(y_2 + x_1),\end{aligned}$$

and so

$$f_0(y_1) - p(y_1 - x_1) \leq p(y_2 + x_1) - f_0(y_2).$$

It follows that

$$\sup\{ f_0(y) - p(y - x_1) : y \in Y \} \leq \inf\{ p(y + x_1) - f_0(y) : y \in Y \},$$

so there is a real number t_1 such that

$$\sup\{ f_0(y) - p(y - x_1) : y \in Y \} \leq t_1 \leq \inf\{ p(y + x_1) - f_0(y) : y \in Y \}.$$

Let $f_1(y + tx_1) = f_0(y) + t \cdot t_1$ for each y in Y and each real number t. It is easy to check that f_1 is a linear functional on Y_1 whose restriction to Y is f_0. It follows from the definition of t_1 that for each y in Y and each positive t,

$$f_1(y + tx_1) = t\big(f_0(t^{-1}y) + t_1\big) \leq tp(t^{-1}y + x_1) = p(y + tx_1)$$

and

$$f_1(y - tx_1) = t\big(f_0(t^{-1}y) - t_1\big) \leq tp(t^{-1}y - x_1) = p(y - tx_1),$$

so $f_1(x) \leq p(x)$ whenever $x \in Y_1$.

The second step of the proof is to show that, in effect, the first step can be repeated until a linear functional on all of X is obtained that is dominated by p and whose restriction to Y is f_0. Let $\mathfrak{A}$ be the collection of all linear functionals g such that the domain of g is a subspace of X that includes Y, the restriction of g to Y is f_0, and g is dominated by p. Define a preorder $\preceq$ on $\mathfrak{A}$ by declaring that $g_1 \preceq g_2$ whenever g_1 is the restriction of g_2 to a subspace of the domain of g_2. It is easy to see that each nonempty chain $\mathfrak{C}$ in $\mathfrak{A}$ has an upper bound in $\mathfrak{A}$; consider the linear functional whose domain is the union Z of the domains of the members of $\mathfrak{C}$ and which agrees at each point z of Z with every member of $\mathfrak{C}$ that is defined at z. Of course, the empty chain has f_0 as an upper bound. By Zorn's lemma, the preordered set $\mathfrak{A}$ has a maximal element f. The domain of f must be all of X, for if it were not then the argument of the first step of this proof, applied to f and its domain instead of to f_0 and Y, would yield an f_1 in $\mathfrak{A}$ such that $f \preceq f_1$ but $f_1 \npreceq f$. This f does all that is required of it. ■

The next result is sometimes called the *analytic form of the Hahn-Banach theorem*. It was first proved by Hahn [99] in 1927 for real normed spaces. Banach [11] independently published the same result with the same proof in 1929, but later became aware of Hahn's earlier paper and acknowledged Hahn's priority. Bohnenblust and Sobczyk [32] gave the extension to complex normed spaces in 1938. In the same year, Soukhomlinoff [227] independently published the same result, and showed that it even holds for the generalization of normed spaces for which the scalars are quaternions.

1.9.6 The Normed Space Version of the Hahn-Banach Extension Theorem. (H. Hahn, 1927, and others; see above). *Suppose that f_0 is a bounded linear functional on a subspace Y of a normed space X. Then there is a bounded linear functional f on all of X such that $\|f\| = \|f_0\|$ and the restriction of f to Y is f_0. That is, the functional f_0 can be extended to a bounded linear functional on X having the same norm.*

PROOF. Suppose that X is a real normed space. Let $p(x) = \|f_0\|\,\|x\|$ for each x in X. Then p is a sublinear functional on X and $f_0(y) \le p(y)$ for each y in Y. By the vector space version of the Hahn-Banach extension theorem, there is a linear functional f on X such that f agrees with f_0 on Y and $f(x) \le \|f_0\|\,\|x\|$ for each x in X, and in fact $|f(x)| \le \|f_0\|\,\|x\|$ for each x in X since

$$-f(x) = f(-x) \le \|f_0\|\,\|-x\| = \|f_0\|\,\|x\|.$$

It follows that f is bounded and $\|f\| \le \|f_0\|$. Since

$$\|f\| = \sup\{\,|f(x)| : x \in B_X\,\} \ge \sup\{\,|f(y)| : y \in B_Y\,\} = \|f_0\|,$$

the linear functionals f_0 and f have the same norm.

Now suppose that X is a complex normed space. Let u_0 be the real part of f_0. By the argument just given for real normed spaces, there is a bounded real-linear functional u on X that agrees with u_0 on Y and has the same norm as u_0. Let f be the unique complex-linear functional on X with real part u. Then the restriction of f to Y is the unique complex-linear functional on Y with real part u_0, that is, the restriction is f_0. Finally, the boundedness of u implies that of f, and $\|f\| = \|u\| = \|u_0\| = \|f_0\|$. ■

The bounded linear functional f guaranteed by the theorem is called a *Hahn-Banach extension* of f_0 to X.

The preceding theorem says nothing about the uniqueness of a Hahn-Banach extension. If f_0 is a bounded linear functional on a dense subspace Y of a normed space X, then Theorem 1.9.1 assures that there is exactly one Hahn-Banach extension of f_0 to all of X. If Y is not dense in X, then f_0 might have more than one Hahn-Banach extension to X, but it also might not. See Exercise 1.105.

1.9.7 Corollary. *Let Y be a closed subspace of a normed space X. Suppose that $x \in X \setminus Y$. Then there is a bounded linear functional f on X such that $\|f\| = 1$, $f(x) = d(x,Y)$, and $Y \subseteq \ker(f)$.*

PROOF. Let $f_0(y + \alpha x) = \alpha \cdot d(x,Y)$ for each y in Y and each scalar α. Then f_0 is a linear functional on $Y + \langle\{x\}\rangle$ such that $f_0(x) = d(x,Y)$ and $f_0(y) = 0$ for each y in Y. Whenever $y \in Y$ and $\alpha \ne 0$,

$$|f_0(y + \alpha x)| = |\alpha| \cdot d(x,Y) \le |\alpha|\,\|x - (-\alpha^{-1}y)\| = \|y + \alpha x\|,$$

so f_0 is bounded and $\|f_0\| \le 1$. Also,

$$\|f_0\|\,\|x-y\| \ge |f_0(x-y)| = d(x,Y)$$

whenever $y \in Y$, and so

$$\|f_0\| \cdot d(x,Y) = \|f_0\| \cdot \inf\{\, \|x-y\| : y \in Y \,\} \ge d(x,Y).$$

Since $d(x,Y) > 0$, it follows that $\|f_0\| \ge 1$, and so $\|f_0\| = 1$. To finish, let f be any Hahn-Banach extension of f_0 to X. ■

Letting $Y = \{0\}$ in the preceding corollary yields the first of the following two results. The second then follows from the first by replacing x by $x - y$.

1.9.8 Corollary. *If x is a nonzero element of a normed space X, then there is a bounded linear functional f on X such that $\|f\| = 1$ and $f(x) = \|x\|$.*

1.9.9 Corollary. *If x and y are different elements of a normed space X, then there is a bounded linear functional f on X such that $f(x) \ne f(y)$.*

The historical roots of the Hahn-Banach extension theorems can be found in efforts to obtain simultaneous solutions for systems of linear equations. To see the connection, consider what must be done to solve a system of m linear equations in n unknowns. Given a collection of m elements $(\alpha_1^{(1)}, \dots, \alpha_n^{(1)}), \dots, (\alpha_1^{(m)}, \dots, \alpha_n^{(m)})$ of $\mathbb{F}^n$ and a collection of m scalars $c_1, \dots, c_m$, it is necessary to find an element $(\beta_1, \dots, \beta_n)$ of $\mathbb{F}^n$ such that $\alpha_1^{(j)}\beta_1 + \cdots + \alpha_n^{(j)}\beta_n = c_j$ when $j = 1, \dots, m$. It is a standard fact from linear algebra that there is a one-to-one correspondence between $(\mathbb{F}^n)^\#$ and $\mathbb{F}^n$ such that for each f in $(\mathbb{F}^n)^\#$ and the corresponding element $(\beta_1, \dots, \beta_n)$ of $\mathbb{F}^n$,

$$f(\alpha_1, \dots, \alpha_n) = \alpha_1\beta_1 + \cdots + \alpha_n\beta_n$$

whenever $(\alpha_1, \dots, \alpha_n) \in \mathbb{F}^n$. Let $x_j = (\alpha_1^{(j)}, \dots, \alpha_n^{(j)})$ when $j = 1, \dots, m$. Then the problem of finding a solution to the given system becomes that of finding a linear functional f on $\mathbb{F}^n$ such that $f(x_j) = c_j$ for each j.

Now suppose that Y is a subspace of a normed space X and that f_0 is a bounded linear functional on Y. For each y in Y, let $c_y = f_0(y)$. The normed space version of the Hahn-Banach extension theorem says that there is a bounded linear functional f on X with the same norm as f_0 such that $f(y) = c_y$ for each y in Y. In a sense, a system of linear equations has been solved. Of course, this system is infinite if $Y \ne \{0\}$.

It is not surprising that the Hahn-Banach extension theorems have important applications to solving systems of linear equations. The following one is due to Hahn himself.

1.9.10 Theorem. (H. Hahn, 1927 [99]). *Suppose that X is a normed space. Let A be a nonempty subset of X and let $\{ c_x : x \in A \}$ be a corresponding collection of scalars. Then the following are equivalent.*

(a) *There is a bounded linear functional f on X such that $f(x) = c_x$ for each x in A.*

(b) *There is a nonnegative real number M such that*

$$|\alpha_1 c_{x_1} + \cdots + \alpha_n c_{x_n}| \le M \, \|\alpha_1 x_1 + \cdots + \alpha_n x_n\|$$

for each linear combination $\alpha_1 x_1 + \cdots + \alpha_n x_n$ of elements of A, that is, for each element of $\langle A \rangle$.

If (b) *holds, then f can be chosen in* (a) *so that $\|f\| \le M$.*

Proof. Suppose that (b) holds. For each member x of $\langle A \rangle$, first express x as a linear combination $\alpha_1 x_1 + \cdots + \alpha_n x_n$ of members of A, then let

$$f_0(x) = f_0(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \alpha_1 c_{x_1} + \cdots + \alpha_n c_{x_n}.$$

If a member of $\langle A \rangle$ is expressed in two different ways $\alpha_1 x_1 + \cdots + \alpha_n x_n$ and $\beta_1 y_1 + \cdots + \beta_m y_m$ as a linear combination of members of A, then

$$\left| \sum_{j=1}^{n} \alpha_j c_{x_j} - \sum_{j=1}^{m} \beta_j c_{y_j} \right| \le M \left\| \sum_{j=1}^{n} \alpha_j x_j - \sum_{j=1}^{m} \beta_j y_j \right\| = 0,$$

so $\alpha_1 c_{x_1} + \cdots + \alpha_n c_{x_n} = \beta_1 c_{y_1} + \cdots + \beta_m c_{y_m}$. Therefore f_0 is an unambiguously defined scalar-valued function on $\langle A \rangle$. It is also clear that f_0 is linear. The inequality in (b) assures that f_0 is bounded and that $\|f_0\| \le M$, so every Hahn-Banach extension f of f_0 to X satisfies (a) and has norm no more than M.

Conversely, suppose that (a) holds. Then for each linear combination $\alpha_1 x_1 + \cdots + \alpha_n x_n$ of elements of A,

$$\begin{aligned} |\alpha_1 c_{x_1} + \cdots + \alpha_n c_{x_n}| &= |f(\alpha_1 x_1 + \cdots + \alpha_n x_n)| \\ &\le \|f\| \, \|\alpha_1 x_1 + \cdots + \alpha_n x_n\|. \end{aligned}$$

Let $M = \|f\|$ to obtain (b). ■

It is possible to exchange the roles of the elements of X and of the bounded linear functionals on X in the preceding theorem, though only finite collections of bounded linear functionals can be considered and the norm of the object found in (a) must be enlarged a bit. The resulting fact, called *Helly's theorem,* can be treated as a straightforward consequence of the normed space version of the Hahn-Banach extension theorem, though it actually predates it by a few years. The following useful result from linear algebra is needed for the proof of Helly's theorem given below.

1.9.11 Lemma. *Suppose that f and $f_1, \dots, f_n$ are linear functionals on the same vector space. Then f is a linear combination of $f_1, \dots, f_n$ if and only if* $\ker(f_1) \cap \cdots \cap \ker(f_n) \subseteq \ker(f)$.

PROOF. The "only if" portion is obvious. For the converse, proceed by induction on n. Let P_k be the proposition that if f and $f_1, \dots, f_k$ are linear functionals on the same vector space and $\bigcap_{j=1}^k \ker(f_j) \subseteq \ker(f)$, then f must be a linear combination of $f_1, \dots, f_k$.

To prove P_1, suppose that f and f_1 are linear functionals on a vector space V such that $\ker(f_1) \subseteq \ker(f)$. The desired conclusion is obvious if f is the zero functional, so it can be assumed that f is nonzero and thus that f_1 is also nonzero. The following standard fact from linear algebra is now needed: If g is a nonzero linear functional on V, then $\ker(g)$ is a maximal proper subspace of V; see, for example, [105]. It follows that $\ker(f_1) = \ker(f)$. Let x be a member of $V \setminus \ker(f)$ and let $h = f - \big(f(x)/f_1(x)\big)f_1$. Then the linear functional h is zero on $\ker(f)$ and is zero at x, and so is zero on V since $V = \langle \{x\} \cup \ker(f) \rangle$. It follows that $f = \big(f(x)/f_1(x)\big)f_1$, which proves P_1.

Suppose that $k \geq 2$ and that P_j is true when $j = 1, \dots, k-1$. Let f and $f_1, \dots, f_k$ satisfy the hypotheses of P_k. If g and $g_1, \dots, g_{k-1}$ are the restrictions of f and $f_1, \dots, f_{k-1}$ respectively to the vector space $\ker(f_k)$, then $\bigcap_{j=1}^{k-1} \ker(g_j) \subseteq \ker(g)$, so by P_{k-1} there are scalars $\alpha_1, \dots, \alpha_{k-1}$ such that $g = \sum_{j=1}^{k-1} \alpha_j g_j$. It follows that $\ker(f_k) \subseteq \ker(f - \sum_{j=1}^{k-1} \alpha_j f_j)$, so by P_1 there is a scalar α_k such that $f - \sum_{j=1}^{k-1} \alpha_j f_j = \alpha_k f_k$. This proves P_k and finishes the induction. ■

1.9.12 Helly's Theorem. (E. Helly, 1921 [102]). *Suppose that X is a normed space. Let $f_1, \dots, f_n$ be a nonempty finite collection of bounded linear functionals on X and let $c_1, \dots, c_n$ be a corresponding collection of scalars. Then the following are equivalent.*

(a) *There is an x_0 in X such that $f_j(x_0) = c_j$ when $j = 1, \dots, n$.*

(b) *There is a nonnegative real number M such that*

$$|\alpha_1 c_1 + \cdots + \alpha_n c_n| \leq M \, \|\alpha_1 f_1 + \cdots + \alpha_n f_n\|$$

for each linear combination $\alpha_1 f_1 + \cdots + \alpha_n f_n$ of $f_1, \dots, f_n$, that is, for each element of $\langle \{f_1, \dots, f_n\} \rangle$.

If (b) *holds and $\epsilon > 0$, then x_0 can be chosen in* (a) *so that $\|x_0\| \leq M + \epsilon$.*

PROOF. If (a) holds, then whenever $\alpha_1, \dots, \alpha_n \in \mathbb{F}$,

$$\begin{aligned} |\alpha_1 c_1 + \cdots + \alpha_n c_n| &= |(\alpha_1 f_1 + \cdots + \alpha_n f_n)(x_0)| \\ &\leq \|\alpha_1 f_1 + \cdots + \alpha_n f_n\| \, \|x_0\|. \end{aligned}$$

Let $M = \|x_0\|$ to obtain (b).

Conversely, suppose that (b) holds and that $\epsilon > 0$. If each c_j is zero, then both (a) and the remark following (b) hold when $x_0 = 0$, so it can be assumed that some c_j is nonzero. It follows from the inequality in (b) that some f_j is nonzero, so it can be assumed, after rearranging the f_j's if necessary, that there is an integer m for which $f_1, \ldots, f_m$ is a maximal linearly independent subcollection of $f_1, \ldots, f_n$. Thus, for each k such that $k = 1, \ldots, n$ there are scalars $\alpha_1^{(k)}, \ldots, \alpha_m^{(k)}$ such that $f_k = \sum_{j=1}^m \alpha_j^{(k)} f_j$.

Suppose that it were known that (b) implies (a) and the remark following (b) under the additional assumption that the linear functionals $f_1, \ldots, f_n$ are linearly independent. In the current situation in which only $f_1, \ldots, f_m$ are known to be linearly independent, it would still follow that there is an x_0 in X such that $\|x_0\| \leq M + \epsilon$ and $f_j(x_0) = c_j$ when $j = 1, \ldots, m$, which implies that $f_k(x_0) = \sum_{j=1}^m \alpha_j^{(k)} c_j$ for each k. It would then follow that

$$|f_k(x_0) - c_k| = \left|\sum_{j=1}^{m} \alpha_j^{(k)} c_j - c_k\right| \leq M \left\|\sum_{j=1}^{m} \alpha_j^{(k)} f_j - f_k\right\| = 0,$$

that is, that $f_k(x_0) = c_k$, when $k = 1, \ldots, n$. Thus, it can be assumed that $f_1, \ldots, f_n$ are linearly independent.

For each x in X, let $T(x) = (f_1 x, \ldots, f_n x)$. Then T is a linear operator from X into $\mathbb{F}^n$. Suppose for the moment that $n \geq 2$. For each k such that $k = 1, \ldots, n$, Lemma 1.9.11 implies that $\bigcap_{j \neq k} \ker(f_j) \not\subseteq \ker(f_k)$, so there is a y_k in X such that $f_k(y_k) = 1$ and $f_j(y_k) = 0$ when $j \neq k$. It follows that $T(X)$ includes the standard basis for $\mathbb{F}^n$, so T maps X onto $\mathbb{F}^n$. If $n = 1$, then T obviously maps X onto $\mathbb{F}^n$. In any case, there is a y_0 in X such that $(f_1 y_0, \ldots, f_n y_0) = (c_1, \ldots, c_n)$. Since $y_0 \notin \bigcap_{j=1}^n \ker(f_j)$, it follows from Corollary 1.9.7 that there is a bounded linear functional f on X such that $\|f\| = 1$, $f(y_0) = d\big(y_0, \bigcap_{j=1}^n \ker(f_j)\big)$, and $\bigcap_{j=1}^n \ker(f_j) \subseteq \ker(f)$. Another application of Lemma 1.9.11 yields scalars $\beta_1, \ldots, \beta_n$ such that $f = \sum_{j=1}^n \beta_j f_j$. Therefore

$$d\left(y_0, \bigcap_{j=1}^{n} \ker(f_j)\right) = f(y_0) = \sum_{j=1}^{n} \beta_j c_j \leq M \left\|\sum_{j=1}^{n} \beta_j f_j\right\| = M,$$

so there is a z_0 in $\bigcap_{j=1}^n \ker(f_j)$ such that $\|y_0 - z_0\| \leq M + \epsilon$. It follows that $y_0 - z_0$ is an x_0 that does what is needed in (a) and in the remark following (b). ∎

Though Helly's theorem is a much weakened analog of Theorem 1.9.10, it is the best that can be done without placing additional restrictions on X. See Exercises 1.119 and 1.134.

In addition to the extension theorems, there is another type of Hahn-Banach theorem called a *separation theorem*. The theme of such a theorem is that given two nonempty convex subsets C_1 and C_2 of a vector space X that are not too badly intermingled, it is possible to find a well-behaved nonzero real-linear functional f on X such that

$$\sup\{ f(x) : x \in C_1 \} \leq \inf\{ f(x) : x \in C_2 \},$$

where under certain conditions it is even possible to assert that the inequality is strict. In a sense, the functional f separates C_1 from C_2, since $f(C_1)$ and $f(C_2)$ are disjoint or nearly so. Such separation theorems are based on a very close relationship that exists between nonnegative sublinear functionals and convex absorbing sets.

1.9.13 Definition. Let A be an absorbing subset of a vector space X. For each x in X, let $p_A(x) = \inf\{ t : t > 0,\ x \in tA \}$. Then p_A is the *Minkowski functional* or *gauge functional* of A.

With X, A, and p_A as in the preceding definition, let x be a nonzero member of X. Very roughly speaking, the number $p_A(x)$ is the ratio of the distance from the origin to x and the distance from the origin to the outermost edge of A in the direction of x, where that second distance might be infinite.

1.9.14 Proposition. *Suppose that X is a vector space.*

(a) *Let p_A be the Minkowski functional of an absorbing subset A of X.*
  (1) *The function p_A is finite-valued, nonnegative, and positive-homogeneous, and $A \subseteq \{ x : x \in X,\ p_A(x) \leq 1 \}$.*
  (2) *If A is a convex set, then p_A is a sublinear functional on X, and $\{ x : x \in X,\ p_A(x) < 1 \} \subseteq A$.*
  (3) *If A is both convex and balanced, then p_A is a seminorm on X.*

(b) *Let p be a nonnegative-real-valued positive-homogeneous function on X and let $A_p = \{ x : x \in X,\ p(x) < 1 \}$.*
  (1) *The set A_p is absorbing, and p is the Minkowski functional of A_p.*
  (2) *If p is a sublinear functional, then A_p is a convex set.*
  (3) *If p is a seminorm, then A_p is both convex and balanced.*

PROOF. Let p_A be the Minkowski functional of an absorbing subset A of X. It follows immediately from the definition of an absorbing set that $\{ t : t > 0,\ x \in tA \}$ is nonempty whenever $x \in X$, so p_A is nonnegative-real-valued. If $x \in X$ and $t_0 > 0$, then

$$t_0\{ t : t > 0,\ x \in tA \} = \{ s : s > 0,\ t_0 x \in sA \},$$

so $t_0 p_A(x) = p_A(t_0 x)$. The function p_A is therefore positive-homogeneous. It is clear that $A \subseteq \{ x : x \in X,\ p_A(x) \leq 1 \}$. Now suppose that A is convex. If $x_1, x_2 \in X$, $t_1, t_2 > 0$, $x_1 \in t_1 A$, and $x_2 \in t_2 A$, then $x_1 + x_2 \in t_1 A + t_2 A = (t_1 + t_2)A$, so $p_A(x_1 + x_2) \le t_1 + t_2$. Taking appropriate infima shows that $p_A(x_1 + x_2) \le p_A(x_1) + p_A(x_2)$, so p_A is a sublinear functional. If $p_A(x) < 1$, then there is a t such that $0 < t < 1$ and $t^{-1}x \in A$, so $x = t(t^{-1}x) + (1-t)0 \in A$. Therefore $\{ x : x \in X, p_A(x) < 1 \} \subseteq A$. Suppose next that A is both convex and balanced, that α is a scalar, that $x \in X$, and that $t > 0$. Then $t^{-1}\alpha x \in A$ if and only if $t^{-1}|\alpha| x \in A$, so $\alpha x \in tA$ if and only if $|\alpha| x \in tA$. It follows that $p_A(\alpha x) = p_A(|\alpha| x) = |\alpha| p_A(x)$ (after a moment's thought about the special case in which $\alpha = 0$), so p_A is a seminorm.

Now let p be a nonnegative-real-valued positive-homogeneous function on X and let $A_p = \{ x : x \in X, p(x) < 1 \}$. If $x \in X$ and $t > p(x)$, then $p(t^{-1}x) < 1$, so $x \in tA_p$. The set A_p is therefore absorbing. For each x in X,

$$\begin{aligned} \inf\{ t : t > 0, x \in tA_p \} &= \inf\{ t : t > 0, p(t^{-1}x) < 1 \} \\ &= \inf\{ t : t > 0, p(x) < t \} \\ &= p(x), \end{aligned}$$

so p is the Minkowski functional of A_p. Now suppose that p is a sublinear functional. If $x, y \in A_p$ and $0 < t < 1$, then

$$p\big(tx + (1-t)y\big) \le tp(x) + (1-t)p(y) < 1,$$

so $tx + (1-t)y \in A_p$. Thus, the set A_p is convex. Finally, suppose that p is a seminorm. Then A_p is convex because p is a sublinear functional. If α is a scalar such that $0 < |\alpha| \le 1$, then

$$\alpha A_p = \{ x : x \in X, p(\alpha^{-1}x) < 1 \} = \{ x : x \in X, p(x) < |\alpha| \} \subseteq A_p.$$

Since A_p is absorbing, it is also true that $0A_p = \{0\} \subseteq A_p$, and so A_p is balanced. ■

In particular, a real-valued function on a vector space is a nonnegative sublinear functional if and only if it is the Minkowski functional of a convex absorbing subset of the space. (However, it does *not* follow that an absorbing set is convex when its Minkowski functional is sublinear. See Exercise 1.108 (b).)

The following result is an example of how Minkowski functionals and the vector space version of the Hahn-Banach extension theorem are used to obtain separation theorems. Far better and more general results of this type will be obtained in Section 2.2, but this proposition will do all of the separating that needs to be done in this first chapter.

1.9.15 Proposition. *If C is a nonempty convex set in a normed space X and x_0 is a point in X such that $d(x_0, C) > 0$, then there is a bounded linear functional f on X such that* $\operatorname{Re} f(x_0) > \sup\{ \operatorname{Re} f(x) : x \in C \}$.

PROOF. Since every bounded real-linear functional on a complex normed space is the real part of a bounded complex-linear functional on the space, it can be assumed that X is a real normed space. Let V be the open ball centered at 0 with radius $2^{-1}d(x_0, C)$ and let y_0 be an element of C. It is easy to check that $-y_0 + (C + V)$ is open, convex, and contains 0 and that $d(x_0 - y_0, -y_0 + (C + V)) = d(x_0, C + V) > 0$. If f is a bounded linear functional on X such that $f(x_0 - y_0) > \sup\{ f(x) : x \in -y_0 + (C+V) \}$, then $f(x_0) > \sup\{ f(x) : x \in C + V \} \geq \sup\{ f(x) : x \in C \}$, so no generality is lost by assuming that C is an open convex set that contains 0. Notice that this implies that C is absorbing.

Let p_C be the Minkowski functional of C, a sublinear functional on X. Since $d(x_0, C) > 0$, there is an s_0 such that $0 < s_0 < 1$ and $s_0 x_0 \notin C$. Since $s_0 x_0$ cannot be a convex combination of 0 and another element of C, it follows that $s x_0 \notin C$ whenever $s \geq s_0$, so $x_0 \notin tC$ whenever $0 < t \leq s_0^{-1}$. Thus,

$$p_C(x_0) \geq s_0^{-1} > 1 \geq \sup\{ p_C(x) : x \in C \}.$$

Let $f_0(tx_0) = tp_C(x_0)$ whenever $t \in \mathbb{R}$. Then f_0 is a linear functional on $\langle \{x_0\} \rangle$ that is dominated by p_C, since $f_0(tx_0) = p_C(tx_0)$ when $t > 0$ and $f_0(tx_0) \leq 0 \leq p_C(tx_0)$ when $t \leq 0$. By the vector space version of the Hahn-Banach extension theorem, there is a linear functional f on X that agrees with f_0 on $\langle \{x_0\} \rangle$ and is dominated by p_C on X. Now C includes an open ball U centered at 0, and for each x in U,

$$|f(x)| = \max\{f(x), f(-x)\} \leq \max\{p_C(x), p_C(-x)\} \leq 1.$$

It follows that f is bounded. Finally,

$$f(x_0) = p_C(x_0) > \sup\{ p_C(x) : x \in C \} \geq \sup\{ f(x) : x \in C \},$$

which finishes the proof. ■

Exercises

1.100 Let X be any Banach space with a proper dense subspace Y; for example, the space X could be c_0 and Y could be the subspace of X consisting of the finitely nonzero sequences. Show that the identity operator on Y cannot be extended to a continuous function from X into Y.

1.101 Let X and Y be normed spaces and let T be a bounded finite-rank linear operator from a subspace W of X into Y. Prove that T can be extended to a bounded linear operator on all of X with the same range as T.

1.102 For this exercise, let the scalar field be $\mathbb{R}$. Recall that a sequence of scalars (α_n) is *Cesàro summable* to a scalar α or that α is the $(C, 1)$-*limit* of (α_n) if $\lim_n((\alpha_1 + \cdots + \alpha_n)/n)$ exists and equals α; see, for example, [19].

(a) Prove that there is a bounded linear functional L on ℓ_∞ such that for each member (α_n) of ℓ_∞,

$$L(\alpha_n) \le \limsup_n \frac{\alpha_1 + \cdots + \alpha_n}{n}$$

and $L(\alpha_n)$ is the $(C,1)$-limit of (α_n) if it has one. (The functional L is called a *Banach limit function* on ℓ_∞, and the "limits" that L assigns to the members of ℓ_∞ are said to form a *system of Banach limits.*)

(b) Prove that

$$\begin{aligned}\liminf_n \alpha_n &\le \liminf_n \frac{\alpha_1 + \cdots + \alpha_n}{n} \\ &\le L(\alpha_n) \\ &\le \limsup_n \frac{\alpha_1 + \cdots + \alpha_n}{n} \\ &\le \limsup_n \alpha_n\end{aligned}$$

for each member (α_n) of ℓ_∞. Conclude from this that the notion of Banach limit is a generalization of the notion of $(C,1)$-limit, which is in turn a generalization of the usual notion of limit. Show that each generalization is proper, that is, that no two of the notions are equivalent.

(c) Prove that $L(\alpha_n) = L(\alpha_{k+1}, \alpha_{k+2}, \ldots)$ for each member (α_n) of ℓ_∞ and each positive integer k, that is, that L is *shift-invariant.* Therefore Banach limits of bounded sequences, like regular limits, depend only on the tails of the sequences and not on their leading terms.

(d) Of the three notions of limit discussed in (b), for which ones is it true that each subsequence of a convergent bounded sequence must converge to the limit of the entire sequence?

(e) Show that the linear functional L found in part (a) is not unique, that is, that there is more than one system of Banach limits for ℓ_∞.

1.103 Suppose that (f_n) is a bounded sequence of bounded linear functionals on a real normed space X. Show that there is a bounded linear functional f on X such that $\liminf_n f_n(x) \le f(x) \le \limsup_n f_n(x)$ whenever $x \in X$. Conclude from this that if $\lim_n f_n(x)$ exists whenever $x \in X$, then the formula $f(x) = \lim_n f_n(x)$ defines a bounded linear functional on X. (Compare Corollary 1.6.10.)

1.104 Let Y be a subspace of a normed space X and let $\mathfrak{A}$ be the collection of all bounded linear functionals f on X such that $Y \subseteq \ker(f)$. Prove that $\overline{Y} = \bigcap\{\ker(f) : f \in \mathfrak{A}\}$.

1.105 Let X be the vector space $\mathbb{R}^2$ and let Y be the subspace of X consisting of all (α, β) such that $\beta = 0$. Define $f_0 \colon Y \to \mathbb{R}$ by the formula $f_0(\alpha, 0) = \alpha$. Notice that f_0 is a bounded linear functional on Y with respect to any norm given to X.

(a) Suppose that X is given the Euclidean norm. Show that f_0 has a unique Hahn-Banach extension to X.

(b) Suppose that X is given the norm that makes it into ℓ_1^2. Show that f_0 has an infinite number of different Hahn-Banach extensions to X.

1.106 Let X be a normed space and let W be a finite-dimensional subspace of $B(X,\mathbb{F})$. Suppose that F is a linear functional on $B(X,\mathbb{F})$. Show that there is an x_0 in X such that $F(w) = w(x_0)$ for every w in W. If F is bounded and $\epsilon > 0$, show that x_0 can be chosen so that $\|x_0\| \leq \|F\| + \epsilon$.

1.107 Which of the properties positive-homogeneity, subadditivity, and sublinearity guarantee that if a real-valued function p on a vector space has that property, then $p(0) = 0$?

1.108 (a) Prove that every norm is the Minkowski functional of the closed unit ball of the corresponding normed space.

(b) Find a nonconvex absorbing subset of some vector space such that the Minkowski functional of the set is a norm.

1.109 Let X be a normed space. If $t \in \mathbb{R}$ and f is a nonzero bounded linear functional on X, then the subset of X defined by the formula

$$H(f,t) = \{\, x : x \in X,\ \operatorname{Re} f(x) \leq t \,\}$$

is called a *closed halfspace*. Show that if $X \neq \{0\}$, then each closed convex subset of X is the intersection of some collection of closed halfspaces. (Don't forget special cases.) What can be said when $X = \{0\}$?

1.110 Suppose that C is a nonempty convex subset of a normed space X and that x_0 is an element of $X \setminus C$ such that $d(x_0, C) = 0$. Proposition 1.9.15 might lead one to conjecture that there must be some nonzero bounded linear functional f on X such that $\operatorname{Re} f(x_0) \geq \sup\{\operatorname{Re} f(x) : x \in C\}$. Find a counterexample. (One can be constructed using a proper dense subspace of a normed space.)

1.10 Dual Spaces

The term *dual space* is often used in linear algebra to refer to the vector space $X^{\#}$ of all linear functionals on a vector space X. In the theory of normed spaces, the vector spaces being studied often have norm topologies on them, and the unbounded linear functionals on these spaces are usually not nearly as important as the bounded ones. For that reason, it is customary in this context to reserve the term *dual space* for the space of all *bounded* linear functionals on a normed space.

1.10.1 Definition. (H. Hahn, 1927 [99]). Let X be a normed space. The *(continuous) dual space* of X or *dual* of X or *conjugate space* of X is the normed space $B(X,\mathbb{F})$ of all bounded linear functionals on X with the operator norm. This space is denoted by X^*.

To avoid confusion, the vector space $X^{\#}$ of all linear functionals on a normed space X is called the *algebraic dual space* of X. Notice that when X is a finite-dimensional normed space, the vector space underlying X^* is just $X^{\#}$, for in this case Theorem 1.4.12 implies that every linear functional on X is bounded. However, infinite-dimensional normed spaces always have linear functionals on them that are unbounded, as was shown in Theorem 1.4.11.

1.10.2 Example. Let (Ω, Σ, μ) be a σ-finite positive measure space. Fix p such that $1 \le p < \infty$ and let L_p denote $L_p(\Omega, \Sigma, \mu)$. Let q be conjugate to p; that is, let $q = \infty$ if $p = 1$ and let q be such that $p^{-1} + q^{-1} = 1$ otherwise. If x^* is a bounded linear functional on L_p, then an argument involving the Radon-Nikodým theorem yields a g in L_q such that

$$x^*(f) = \int_\Omega fg\,d\mu \tag{1.2}$$

whenever $f \in L_p$. Conversely, for each g in L_q, Hölder's inequality assures that (1.2) defines a bounded linear functional x^* on L_p. Moreover, if an x^* in L_p^* and a g in L_q are related by (1.2), then $\|x^*\| = \|g\|_q$. See [202], for example, for the detailed arguments. If $T\colon L_q \to L_p^*$ is defined by the formula $T(g) = x^*$, where x^* is given by (1.2), then it follows that T is an isometric isomorphism from L_q onto L_p^*. Because of this, it is common practice to identify L_q with L_p^* and say that L_q *is the dual of* L_p.

In the preceding example, the requirement that the measure space be σ-finite is not necessary *as long as p is not* 1; it can be shown that L_p^* can still be identified in the above way with L_q. If the measure space is not σ-finite, then it can happen that L_1^* is not L_∞ in the above sense. See pp. 286–288 and 387 of [67] for a discussion of the non-σ-finite case, including some special conditions under which L_1^* must still be L_∞.

There is a natural way to identify L_∞^* with a normed space $\mathrm{ba}(\Omega, \Sigma_1, \mu_1)$ whose underlying set consists of all the bounded finitely additive scalar-valued set functions with domain Σ_1, where $(\Omega, \Sigma_1, \mu_1)$ is a measure space related to (Ω, Σ, μ). See [67] for details.

1.10.3 Example. Suppose that μ is the counting measure on the collection Σ of all subsets of $\mathbb{N}$ and that $1 \le p < \infty$. Let q be conjugate to p. Since ℓ_p is $L_p(\mathbb{N}, \Sigma, \mu)$, it follows from the preceding example that ℓ_p^* can be identified with ℓ_q. If the element x^* of ℓ_p^* corresponds to the element (β_n) of ℓ_q, then the action of x^* on an element (α_n) of ℓ_p is given by the formula

$$x^*(\alpha_n) = \sum_n \alpha_n \beta_n.$$

Though the dual space of ℓ_∞ cannot be characterized so easily as a space of sequences, it can be identified in a natural way with a normed space whose underlying set consists of all the bounded finitely additive scalar-valued set functions on the subsets of $\mathbb{N}$. See [58] for an excellent discussion of ℓ_∞^*.

1.10.4 Example. The dual space of c_0 is ℓ_1 in exactly the same sense that the dual space of ℓ_p is ℓ_q when $1 \le p < \infty$ and q is conjugate to p. That is, there is an isometric isomorphism T from ℓ_1 onto c_0^* such that if $(\beta_n) \in \ell_1$, then $T(\beta_n)$ is the linear functional x^* on c_0 given by the formula

$$x^*(\alpha_n) = \sum_n \alpha_n \beta_n. \tag{1.3}$$

Unlike the preceding example, this is not just a special case of Example 1.10.2 since c_0 is not a Lebesgue space in any natural sense. However, it is not difficult to prove directly.

If $(\beta_n) \in \ell_1$, then $\sum_n |\alpha_n \beta_n| \le \|(\beta_n)\|_1 \, \|(\alpha_n)\|_\infty$ for each (α_n) in c_0, from which it follows that (1.3) defines a member x^* of c_0^* for which $\|x^*\| \le \|(\beta_n)\|_1$. It is then easy to check that the function $T\colon \ell_1 \to c_0^*$ that maps each (β_n) in ℓ_1 to the x^* defined by (1.3) is a linear operator such that $\|T(\beta_n)\| \le \|(\beta_n)\|_1$ whenever $(\beta_n) \in \ell_1$, and this operator is one-to-one since $T(\beta_n) \ne 0$ when $(\beta_n) \ne 0$.

Now suppose that x^* is an arbitrary element of c_0^*. For each standard unit vector e_n of c_0, let $\beta_n = x^* e_n$ and let γ_n be a scalar such that $|\gamma_n| = 1$ and $|\beta_n| = \gamma_n \beta_n$. Then

$$\sum_{n=1}^m |\beta_n| = \sum_{n=1}^m \gamma_n x^* e_n = x^*\left(\sum_{n=1}^m \gamma_n e_n\right) \le \|x^*\| \left\|\sum_{n=1}^m \gamma_n e_n\right\|_\infty = \|x^*\|$$

for each m. The sequence (β_n) is therefore in ℓ_1, and $\|(\beta_n)\|_1 \le \|x^*\|$. For each member (α_n) of c_0,

$$x^*(\alpha_n) = x^*\left(\sum_n \alpha_n e_n\right) = \sum_n \alpha_n x^* e_n = \sum_n \alpha_n \beta_n,$$

so x^* and (β_n) satisfy (1.3) whenever $(\alpha_n) \in c_0$. Since $T(\beta_n) = x^*$, the operator T maps ℓ_1 onto c_0^*.

Finally, suppose that (β_n) is an arbitrary member of ℓ_1. Let $x^* = T(\beta_n)$. By the argument of the preceding paragraph, there is a member (β_n') of ℓ_1 such that $T(\beta_n') = x^*$ and $\|(\beta_n')\|_1 \le \|x^*\|$. Since T is one-to-one, it follows that $(\beta_n') = (\beta_n)$ and therefore that $\|(\beta_n)\|_1 \le \|T(\beta_n)\|$. Since it has already been shown that $\|T(\beta_n)\| \le \|(\beta_n)\|_1$, the operator T is an isometric isomorphism, which finishes the proof.

1.10.5 Example. Suppose that n is a positive integer and that $1 \le p < \infty$. Let q be conjugate to p. If μ is the counting measure on the collection Σ of all subsets of $\{1, \dots, n\}$, then ℓ_p^n is $L_p(\{1, \dots, n\}, \Sigma, \mu)$, and so $(\ell_p^n)^*$ can be identified with ℓ_q^n by the method of Example 1.10.2. Notice that if

$(\beta_1, \ldots, \beta_n)$ is in ℓ_q^n and x^* is the linear functional that it represents, then the action of x^* on ℓ_p^n is given by the formula

$$x^*(\alpha_1, \ldots, \alpha_n) = \sum_{j=1}^{n} \alpha_j \beta_j.$$

The dual space of ℓ_∞^n can be identified with ℓ_1^n in the same way. A proof of this, using an argument similar to that of Example 1.10.4, is given in Appendix C; see Theorem C.13. This will also follow from some of the results of the next section, as will be shown in the comments following Theorem 1.11.9.

For the sake of completeness, it should be noted that if $1 \leq p \leq \infty$ and q is conjugate to p, then $(\ell_p^0)^*$ and ℓ_q^0 are both zero-dimensional, so $(\ell_p^0)^*$ can be identified with ℓ_q^0 in a very obvious and trivial way.

1.10.6 Example. Let K be a compact Hausdorff space. The Riesz representation theorem says that $C(K)^*$ is isometrically isomorphic to $\text{rca}(K)$. If x^* is a bounded linear functional on $C(K)$ and μ is the measure in $\text{rca}(K)$ identified with x^*, then the action of x^* on $C(K)$ is given by the formula

$$x^*(f) = \int_K f \, d\mu.$$

See [67] or [202] for details. It turns out that $\text{rca}(K)$ and $C(K)^*$ both have natural partial orders, and that the isometric isomorphism of the Riesz representation theorem is *order preserving* in the sense that if x_1^* and x_2^* are elements of $C(K)^*$ and μ_1 and μ_2 are the respective elements of $\text{rca}(K)$ with which x_1^* and x_2^* are identified, then $x_1^* \preceq x_2^*$ if and only if $\mu_1 \preceq \mu_2$. See Exercise 1.121.

Recall that if X is a normed space and Y is a Banach space, then $B(X, Y)$ is a Banach space; see Theorem 1.4.8. Letting $Y = \mathbb{F}$ in this result produces the first theorem of this section.

1.10.7 Theorem. *If X is a normed space, then X^* is a Banach space.*

Because of this, one way to show that a normed space is a Banach space is to show that it is the dual of some other normed space. For example, if K is a compact Hausdorff space, then $\text{rca}(K)$ is a Banach space since it is isometrically isomorphic to the dual of $C(K)$. See the remarks following Example 1.2.11.

One of the main reasons for proving the Hahn-Banach extension theorems before discussing duality is to obtain Corollary 1.9.9, which says that if x and y are different elements of a normed space X, then there is an x^*

in X^* such that $x^*x \neq x^*y$. That is, the dual space of X is large enough to separate the points of X. Of course, this implies that $X^* \neq \{0\}$ when $X \neq \{0\}$. As will be seen in Sections 2.2 and 2.3, there are large vector spaces with metric topologies possessing many of the properties of norm topologies, even though the spaces have no continuous linear functionals on them besides the zero functional.

The other corollaries of the normed space version of the Hahn-Banach extension theorem also have applications to the theory of dual spaces, as is shown by the proof of the following result.

1.10.8 Theorem. *A normed space is finite-dimensional if and only if its dual space is finite-dimensional.*

PROOF. Let X be a normed space. If X is finite-dimensional, then it is a standard fact from linear algebra that $X^\#$, which in this case is the same as X^* by Theorem 1.4.12, has the same finite dimension as X.

Now suppose instead that X is infinite-dimensional. Let (x_n) be a linearly independent sequence in X. For each nonnegative integer n, let $F_n = \langle\{ x_j : j \leq n \}\rangle$; notice in particular that $F_0 = \{0\}$. Then for each n, the subspace F_n of X is finite-dimensional and therefore closed, and F_n does not contain x_{n+1}. By Corollary 1.9.7, there is a sequence (x_n^*) of elements of X^* such that $x_n^*x_j = 0$ when $j < n$, but $x_n^*x_n$ is nonzero for each n. If $\sum_{j=1}^k \alpha_j x_j^*$ is a linear combination of the terms of (x_n^*) that equals 0, then applying this linear combination to $x_1, \ldots, x_k$ in that order shows that $\alpha_1 = \cdots = \alpha_k = 0$, from which it follows that (x_n^*) is a linearly independent sequence in X^*. The space X^* is therefore infinite-dimensional. ■

Corollary 1.9.8 is another consequence of the Hahn-Banach extension theorems that is useful in the study of dual spaces. The proof of the next theorem is based on it.

1.10.9 Theorem. *Suppose that x is an element of a normed space X. Then*

$$\|x\| = \sup\{ |x^*x| : x^* \in B_{X^*} \}.$$

Furthermore, this supremum is attained at some point of B_{X^*}.

PROOF. If $x = 0$, then the formula for $\|x\|$ is trivially true and the supremum is attained at every point of B_{X^*}. It may therefore be assumed that $x \neq 0$. Since $|x^*x| \leq \|x^*\|\,\|x\| \leq \|x\|$ whenever $x^* \in B_{X^*}$, it follows that $\|x\| \geq \sup\{ |x^*x| : x^* \in B_{X^*} \}$. By Corollary 1.9.8, there is an x_0^* in X^* such that $\|x_0^*\| = 1$ and $x_0^*x = \|x\|$, and so $\|x\| = \sup\{ |x^*x| : x^* \in B_{X^*} \}$, with the supremum being attained at x_0^*. ■

The fact that the supremum is attained in the formula of the preceding theorem leads naturally to the question of when the supremum is attained in the formula for the norm of a bounded linear functional x^* on a normed space X:

$$\|x^*\| = \sup\{\, |x^*x| : x \in B_X \,\}.$$

If X is finite-dimensional, then the compactness of B_X and the continuity of $|x^*|$ together insure that this supremum is attained, and so it is said that all the elements of X^* are *norm-attaining* functionals. If X is not finite-dimensional, then X^* still contains many norm-attaining functionals for the following reason. If $x \in S_X$, then Corollary 1.9.8 produces an x^* in X^* such that $|x^*x| = \|x^*\| = 1$, so this x^* is norm-attaining. It may be, however, that X^* contains functionals that are not norm-attaining. James's theorem, proved in Section 1.13, gives a very useful characterization of the Banach spaces X such that each element of X^* is a norm-attaining functional.

1.10.10 Example. Let x^* be the element of c_0^* represented by the element (2^{-n}) of ℓ_1. Then $\|x^*\| = 1$. If (α_n) is any element of B_{c_0}, then

$$|x^*(\alpha_n)| = \Big|\sum_n 2^{-n}\alpha_n\Big| \le \sum_n 2^{-n}|\alpha_n| < \sum_n 2^{-n} = 1,$$

so x^* is not norm-attaining. See also Exercise 1.114.

Suppose that X and Y are normed spaces and that $y^* \in Y^*$, $x \in X$, $T \in B(X,Y)$, and $S \in B(Y^*,X^*)$. The statement $y^*Tx = Sy^*x$ can have only one meaning, but that meaning might take a few moments to unravel since the order in which the operations are to be performed differs on the two sides of the equation. The use of parentheses helps, but can still lead to expressions that are visually confusing. For this reason, the notation $\langle x, f\rangle$ is sometimes used for fx when f is a linear functional on X and $x \in X$. The statement $y^*Tx = Sy^*x$ then becomes $\langle Tx, y^*\rangle = \langle x, Sy^*\rangle$, which is easier to grasp. This device is useful in the following result.

1.10.11 Proposition. *Let X and Y be normed spaces.*

(a) *If $T \in L(X,Y)$, then T is bounded if and only if*

$$\sup\{\, |\langle Tx, y^*\rangle| : x \in B_X,\ y^* \in B_{Y^*} \,\}$$

is finite. If T is bounded, then its norm equals this supremum.

(b) *If $T \in L(X,Y^*)$, then T is bounded if and only if*

$$\sup\{\, |\langle y, Tx\rangle| : x \in B_X,\ y \in B_Y \,\}$$

is finite. If T is bounded, then its norm equals this supremum.

PROOF. For (a), suppose that $T \in L(X,Y)$. Theorem 1.10.9 implies that $\|Tx\| = \sup\{\, |\langle Tx, y^*\rangle| : y^* \in B_{Y^*} \,\}$ whenever $x \in X$, so

$$\begin{aligned}\sup\{\, \|Tx\| : x \in B_X \,\} &= \sup\{\, \sup\{\, |\langle Tx, y^*\rangle| : y^* \in B_{Y^*} \,\} : x \in B_X \,\} \\ &= \sup\{\, |\langle Tx, y^*\rangle| : x \in B_X,\, y^* \in B_{Y^*} \,\}.\end{aligned}$$

Part (a) follows from this. For (b), suppose that $T \in L(X,Y^*)$. Then $\|Tx\| = \sup\{\, |\langle y, Tx\rangle| : y \in B_Y \,\}$ whenever $x \in X$, so

$$\begin{aligned}\sup\{\, \|Tx\| : x \in B_X \,\} &= \sup\{\, \sup\{\, |\langle y, Tx\rangle| : y \in B_Y \,\} : x \in B_X \,\} \\ &= \sup\{\, |\langle y, Tx\rangle| : x \in B_X,\, y \in B_Y \,\},\end{aligned}$$

from which (b) follows. ∎

If two normed spaces X and Y are isometrically isomorphic, then they are in a sense identical as normed spaces and so should have identical duals. More properly, the spaces X^* and Y^* should also be isometrically isomorphic. A few moments' thought shows how a proof of this might go. Let T be an isometric isomorphism from X onto Y and let $x_T = Tx$ for each x in X; that is, let x_T be the element of Y identified with x. Then each y^* in Y^* can be thought of as a bounded linear functional $y^*_{T^*}$ on X by letting $y^*_{T^*}x = y^*x_T$ for each x in X. The map $T^*\colon Y^* \to X^*$ given by the formula $T^*y^* = y^*_{T^*}$ should be the desired isometric isomorphism from Y^* onto X^*. This idea, and its extension from isometric isomorphisms to isomorphisms, is behind the proof of the next theorem.

1.10.12 Theorem. *Suppose that X and Y are normed spaces such that there is an isomorphism T from X onto Y. Then the map $T^*\colon Y^* \to X^*$ given by the formula $T^*(y^*) = y^*T$, where y^*T is the usual product of y^* and T, is an isomorphism from Y^* onto X^*, and $\|T^*\| = \|T\|$. If T is an isometric isomorphism, then so is T^*.*

PROOF. It is clear that $T^*(y^*) \in X^*$ whenever $y^* \in Y^*$, and equally clear that T^* is linear. Notice that $\langle Tx, y^*\rangle = \langle x, T^*y^*\rangle$ whenever $x \in X$ and $y^* \in Y^*$. By Proposition 1.10.11 (a),

$$\begin{aligned}\|T\| &= \sup\{\, |\langle Tx, y^*\rangle| : x \in B_X,\, y^* \in B_{Y^*} \,\} \\ &= \sup\{\, |\langle x, T^*y^*\rangle| : x \in B_X,\, y^* \in B_{Y^*} \,\},\end{aligned}$$

so Proposition 1.10.11 (b) implies that T^* is bounded and that $\|T^*\| = \|T\|$. If $y^* \in Y^*$ and $T^*y^* = 0$, then $y^*y = \langle T(T^{-1}y), y^*\rangle = \langle T^{-1}y, T^*y^*\rangle = 0$ whenever $y \in Y$, so $y^* = 0$. It follows that T^* is one-to-one. If x^* is a member of X^*, then $x^*x = \langle Tx, x^*T^{-1}\rangle = \langle x, T^*(x^*T^{-1})\rangle$ whenever $x \in X$, so $T^*(x^*T^{-1}) = x^*$. Thus, the operator T^* maps Y^* onto X^*. Since T^* is a one-to-one bounded linear operator from one Banach space onto another, it is an isomorphism by Corollary 1.6.6.

Finally, suppose that T is an isometric isomorphism. Then for each member y^* of Y^*,

$$\begin{aligned}\|T^*y^*\| &= \sup\{\,|\langle x, T^*y^*\rangle| : x \in B_X\,\}\\ &= \sup\{\,|\langle Tx, y^*\rangle| : x \in B_X\,\}\\ &= \sup\{\,|\langle y, y^*\rangle| : y \in B_Y\,\}\\ &= \|y^*\|,\end{aligned}$$

so T^* is also an isometric isomorphism. ■

The map T^* defined in the preceding theorem is called the *adjoint* of T and is very important in the study of linear operators. As will be seen in Section 3.1, every bounded linear operator T from a normed space X into a normed space Y has a bounded linear adjoint T^* defined as in Theorem 1.10.12, and T^* is an isomorphism (respectively, an isometric isomorphism) from Y^* onto X^* if *and only if* T is an isomorphism (respectively, an isometric isomorphism) from X onto Y. However, it is not true that X and Y must even be isomorphic when there is *some* isometric isomorphism from X^* onto Y^*. See Exercises 1.111 and 1.112.

The rest of this section is devoted to finding representations of the dual spaces of subspaces, quotient spaces, and direct sums of normed spaces. The representation of the dual of a direct sum given in the next theorem is a straightforward generalization of the way in which the dual of n-dimensional Euclidean space ℓ_2^n is ℓ_2^n; see Example 1.10.5.

1.10.13 Theorem. *Let $X_1, \dots, X_n$ be normed spaces. Then there is an isometric isomorphism that identifies $(X_1 \oplus \cdots \oplus X_n)^*$ with $X_1^* \oplus \cdots \oplus X_n^*$ such that if the element y^* of $(X_1 \oplus \cdots \oplus X_n)^*$ is identified with the element $(x_1^*, \dots, x_n^*)$ of $X_1^* \oplus \cdots \oplus X_n^*$, then*

$$y^*(x_1, \dots, x_n) = \sum_{j=1}^{n} x_j^* x_j$$

whenever $(x_1, \dots, x_n) \in X_1 \oplus \cdots \oplus X_n$.

Proof. For each member $(x_1^*, \dots, x_n^*)$ of $X_1^* \oplus \cdots \oplus X_n^*$, let $T(x_1^*, \dots, x_n^*)$ be the linear functional on $X_1 \oplus \cdots \oplus X_n$ given by the formula

$$\langle (x_1, \dots, x_n), T(x_1^*, \dots, x_n^*) \rangle = \sum_{j=1}^{n} x_j^* x_j.$$

Then T is the desired isometric isomorphism from $X_1^* \oplus \cdots \oplus X_n^*$ onto $(X_1 \oplus \cdots \oplus X_n)^*$. There is some checking to be done.

If $(x_1^*, \dots, x_n^*) \in X_1^* \oplus \cdots \oplus X_n^*$, then it is clear that $T(x_1^*, \dots, x_n^*)$ really is a linear functional on $X_1 \oplus \cdots \oplus X_n$, and since Cauchy's inequality implies that

$$\left|\langle (x_1,\dots,x_n), T(x_1^*,\dots,x_n^*)\rangle\right| \le \sum_{j=1}^n \|x_j^*\|\,\|x_j\| \le \|(x_1^*,\dots,x_n^*)\|\,\|(x_1,\dots,x_n)\|$$

whenever $(x_1,\dots,x_n) \in X_1 \oplus \cdots \oplus X_n$, the functional $T(x_1^*,\dots,x_n^*)$ is in $(X_1 \oplus \cdots \oplus X_n)^*$ and has norm no more than $\|(x_1^*,\dots,x_n^*)\|$. It follows that T is a linear operator from $X_1^* \oplus \cdots \oplus X_n^*$ into $(X_1 \oplus \cdots \oplus X_n)^*$ such that $\|T(x_1^*,\dots,x_n^*)\| \le \|(x_1^*,\dots,x_n^*)\|$ whenever $(x_1^*,\dots,x_n^*) \in X_1^* \oplus \cdots \oplus X_n^*$.

Now suppose that $y^* \in (X_1 \oplus \cdots \oplus X_n)^*$. For each j such that $j = 1,\dots,n$, it is clear that the function x_j^* on X_j defined by the formula $x_j^*(x) = y^*(0,\dots,0,x,0,\dots,0)$, where x is the j^{th} component of the n-tuple, is a member of X_j^*. Then $T(x_1^*,\dots,x_n^*) = y^*$, which shows that T maps $X_1^* \oplus \cdots \oplus X_n^*$ onto $(X_1 \oplus \cdots \oplus X_n)^*$.

Suppose that $(x_1^*,\dots,x_n^*) \in X_1^* \oplus \cdots \oplus X_n^*$. All that remains is to prove that $\|T(x_1^*,\dots,x_n^*)\| = \|(x_1^*,\dots,x_n^*)\|$, so it can be assumed that $(x_1^*,\dots,x_n^*) \neq 0$. It is enough to prove that

$$\left\|T\left(\|(x_1^*,\dots,x_n^*)\|^{-1}(x_1^*,\dots,x_n^*)\right)\right\| = 1,$$

so it can be assumed that $\|(x_1^*,\dots,x_n^*)\| = 1$ since

$$\left\|\, \|(x_1^*,\dots,x_n^*)\|^{-1}(x_1^*,\dots,x_n^*)\,\right\| = 1.$$

It has already been shown that $\|T(x_1^*,\dots,x_n^*)\| \le \|(x_1^*,\dots,x_n^*)\| = 1$, and so it is enough to show that $\|T(x_1^*,\dots,x_n^*)\| \ge 1$. For each j such that $j = 1,\dots,n$, let $(y_j^{(k)})_{k=1}^\infty$ be a sequence in B_{X_j} such that $\lim_k |x_j^* y_j^{(k)}| = \|x_j^*\|$; because there is for each j and each k a scalar $\alpha_j^{(k)}$ such that $|\alpha_j^{(k)}| = 1$ and $x_j^*(\alpha_j^{(k)} y_j^{(k)}) = |x_j^* y_j^{(k)}|$, it can be assumed that $x_j^* y_j^{(k)} \ge 0$ for each j and each k and so that $\lim_k x_j^* y_j^{(k)} = \|x_j^*\|$ for each j. Let $x_j^{(k)} = \|x_j^*\| y_j^{(k)}$ for each j and each k. Then $\|(x_1^{(k)},\dots,x_n^{(k)})\| \le \|(x_1^*,\dots,x_n^*)\| = 1$ for each k, and so

$$\begin{aligned}
1 &= \sum_{j=1}^n \|x_j^*\|^2 \\
&= \lim_k \sum_{j=1}^n x_j^* x_j^{(k)} \\
&= \lim_k \langle (x_1^{(k)},\dots,x_n^{(k)}), T(x_1^*,\dots,x_n^*)\rangle \\
&\le \|T(x_1^*,\dots,x_n^*)\|,
\end{aligned}$$

which is the inequality needed to finish the proof. ■

As was mentioned in Section 1.8, there are a number of different formulas used to define the norm of a direct sum $X_1 \oplus \cdots \oplus X_n$ of normed spaces that produce norms equivalent to the one given by the formula of Definition 1.8.1. The problem with using one of these other formulas is that Theorem 1.10.13 might not hold as stated, for using that formula to define the norms of both $X_1 \oplus \cdots \oplus X_n$ and $X_1^* \oplus \cdots \oplus X_n^*$ might not result in $(X_1 \oplus \cdots \oplus X_n)^*$ being isometrically isomorphic to $X_1^* \oplus \cdots \oplus X_n^*$ in the sense of the theorem. See Exercise 1.116.

Some facts about annihilators are needed before the duals of subspaces and quotient spaces can be discussed.

1.10.14 Definition. Let X be a normed space and let A and B be subsets of X and X^* respectively. Define $A^\perp$ and ${}^\perp B$ (pronounced "A perp" and "perp B") by the formulas

$$A^\perp = \{\, x^* : x^* \in X^*,\ x^*x = 0 \text{ for each } x \text{ in } A \,\};$$
$${}^\perp B = \{\, x : x \in X,\ x^*x = 0 \text{ for each } x^* \text{ in } B \,\}.$$

Then $A^\perp$ is the *annihilator of A in X^**, while ${}^\perp B$ is the *annihilator of B in X*.

The sets $A^\perp$ and ${}^\perp B$ are usually called the annihilators of A and B respectively, with the qualifying phrases "in X^*" and "in X" dropped. In theory, this could cause confusion when referring to annihilators of subsets of dual spaces. Since the dual X^* of a normed space X is itself a normed space, both ${}^\perp B$ and $B^\perp$ are defined for each subset B of X^*, and both have the right to be called the annihilator of B. In practice, the context usually prevents confusion. Where a misunderstanding might occur, the space in which the annihilator is being taken should be made explicit, either by adding the qualifying phrase or by using the left-hand or right-hand "perp" notation.

1.10.15 Proposition. *Let X be a normed space and let A and B be subsets of X and X^* respectively.*

(a) *The sets $A^\perp$ and ${}^\perp B$ are closed subspaces of X^* and X respectively.*

(b) ${}^\perp(A^\perp) = [A]$.

(c) *If A is a subspace of X, then ${}^\perp(A^\perp) = \overline{A}$.*

PROOF. Since ${}^\perp B = \bigcap\{\ker(x^*) : x^* \in B\}$, it follows that ${}^\perp B$ is a closed subspace of X. Now $0 \in A^\perp$, and if α is a scalar and $x^*, y^* \in A^\perp$, then $\alpha x^* + y^* \in A^\perp$, so $A^\perp$ is a subspace of X^*. Furthermore, if (z_n^*) is a sequence in $A^\perp$ converging to some z^* in X^*, then $z^* \in A^\perp$, so the subspace $A^\perp$ is closed. This proves (a).

For (b), first notice that ${}^\perp(A^\perp)$ is a closed subspace of X that includes A, so $[A] \subseteq {}^\perp(A^\perp)$. Now suppose that $x_0 \in X \setminus [A]$. By Corollary 1.9.7, there is an x_0^* in X^* such that $x_0^*x_0 \neq 0$ and $[A] \subseteq \ker(x_0^*)$. It follows that $x_0 \notin {}^\perp(A^\perp)$, since $x_0^* \in A^\perp$ and $x_0^*x_0 \neq 0$, so $[A] \supseteq {}^\perp(A^\perp)$. This finishes the proof of (b). Since $[A] = \overline{\langle A \rangle}$ whenever $A \subseteq X$, part (c) follows immediately. ■

See Exercise 1.120 for some observations about $({}^\perp B)^\perp$ when B is a subset of the dual space of a normed space. Characterizations of $({}^\perp B)^\perp$ analogous to those obtained in the above proposition for ${}^\perp(A^\perp)$ will be given in Proposition 2.6.6.

Suppose that M is a subspace of a normed space X. The restriction of an x^* in X^* to M always results in an element of M^*. Moreover, if $m^* \in M^*$, then the normed space version of the Hahn-Banach extension theorem allows m^* to be extended to an element $x_{m^*}^*$ of X^*, and the restriction of $x_{m^*}^*$ to M is m^*. Thus, the elements of M^* are just the restrictions of the elements of X^* to M, so in one sense M^* can be identified with X^*. Notice however that different elements of X^* might agree on M, so elements of M^* should really be considered to be equivalence classes of elements of X^*, with two elements of X^* considered equivalent if they agree on M. A moments' thought shows that the resulting equivalence classes are exactly the cosets $x^* + M^\perp$ that form the elements of $X^*/M^\perp$. The following result is now not too surprising.

1.10.16 Theorem. *Let M be a subspace of a normed space X. Then there is an isometric isomorphism that identifies M^* with $X^*/M^\perp$ such that if an element of M^* is identified with the element $x^* + M^\perp$ of $X^*/M^\perp$, then the action of $x^* + M^\perp$ on M is given by the formula $(x^* + M^\perp)(m) = x^*m$.*

PROOF. Let $\big(T(x^* + M^\perp)\big)(m) = x^*m$ whenever $x^* + M^\perp \in X^*/M^\perp$ and $m \in M$; that is, let $T\colon X^*/M^\perp \to M^*$ be the function that maps each member $x^* + M^\perp$ of $X^*/M^\perp$ to the restriction of x^* to M. The goal is to show that T is an isometric isomorphism from $X^*/M^\perp$ onto M^*. Since two elements $x_1^* + M^\perp$ and $x_2^* + M^\perp$ of $X^*/M^\perp$ are equal if and only if x_1^* and x_2^* agree on M, there is no ambiguity in the definition of T. It is clear that T is linear. If $m^* \in M^*$ and $x_{m^*}^*$ is a Hahn-Banach extension of m^* to X, then $T(x_{m^*}^* + M^\perp) = m^*$, so T maps $X^*/M^\perp$ onto M^*.

Suppose that $x^* + M^\perp \in X^*/M^\perp$ and that $m^* = T(x^* + M^\perp)$. Since m^* has a Hahn-Banach extension $x_{m^*}^*$ to X and $x_{m^*}^* + M^\perp = x^* + M^\perp$, it can be assumed that $\|x^*\| = \|m^*\|$. If $y^* \in M^\perp$, then m^* and $x^* + y^*$ agree on M, so

$$\begin{aligned} \|m^*\| &= \sup\{\, |(x^* + y^*)(m)| : m \in B_M \,\} \\ &\leq \sup\{\, |(x^* + y^*)(x)| : x \in B_X \,\} \\ &= \|x^* + y^*\|. \end{aligned}$$

It follows that

$$\|m^*\| \leq \inf\{\, \|x^* + y^*\| : y^* \in M^\perp \,\} = \|x^* + M^\perp\| \leq \|x^*\| = \|m^*\|,$$

so $\|T(x^* + M^\perp)\| = \|m^*\| = \|x^* + M^\perp\|$. Thus, the map T is an isometric isomorphism from $X^*/M^\perp$ onto M^*. ■

Now suppose that M is a *closed* subspace of a normed space X. Since M^* can be identified in a natural way with a normed space derived from X^*, namely $X^*/M^\perp$, it is reasonable to ask if $(X/M)^*$ can also be identified with some normed space related to X^*. A logical first step in finding such an identification is to see how the elements of X^* operate on those of X/M through the formula $x^*(x + M) = x^*x$. However, this formula does not always define $x^*(x+M)$ uniquely. If x_1 and x_2 are different elements of X such that $x_1 - x_2 \in M$, and x^* is an element of X^* that separates x_1 from x_2 in the sense of Corollary 1.9.9, then $x_1 + M = x_2 + M$ but $x^*x_1 \neq x^*x_2$. A bit of reflection reveals that the formula $x^*(x + M) = x^*x$ uniquely defines $x^*(x+M)$ for all elements $x+M$ of X/M precisely when $x^* \in M^\perp$. This suggests the following theorem, which is essentially just a corollary of Theorem 1.7.13.

1.10.17 Theorem. *Let M be a closed subspace of a normed space X. Then there is an isometric isomorphism that identifies $(X/M)^*$ with $M^\perp$ such that if an element of $(X/M)^*$ is identified with the element x^* of $M^\perp$, then the action of x^* on X/M is given by the formula $x^*(x + M) = x^*x$.*

PROOF. Let $\pi\colon X \to X/M$ be the quotient map and let $T(y^*) = y^*\pi$ for each y^* in $(X/M)^*$. It is clear that T is a linear operator from $(X/M)^*$ into $M^\perp$. If $x^* \in M^\perp$, then $M \subseteq \ker(x^*)$, so Theorem 1.7.13 yields a unique y^* in $(X/M)^*$ such that $x^* = y^*\pi$, and the same theorem guarantees that $\|y^*\| = \|x^*\|$. This is just another way to say that T is one-to-one and onto $M^\perp$ and that $\|y^*\| = \|Ty^*\|$ for each y^* in $(X/M)^*$, so T is an isometric isomorphism from $(X/M)^*$ onto $M^\perp$. Notice that if $y^* \in (X/M)^*$ and $x^* = Ty^*$, then $y^*(x+M) = (Ty^*)(x) = x^*x$ whenever $x+M \in X/M$, as required. ■

Exercises

1.111 Let c be the Banach space of all convergent sequences of scalars defined in Exercise 1.25. Prove that c^* is isometrically isomorphic to ℓ_1. Notice from Exercise 1.49 that c and c_0 are not isometrically isomorphic, even though their dual spaces are.

1.112 Let Y be a dense subspace of a normed space X. Prove that X^* and Y^* are isometrically isomorphic. Use this to give an example of two normed spaces X and Y such that X^* and Y^* are isometrically isomorphic, even though X and Y are not even isomorphic.

1.113 Suppose that X is a complex normed space and that $x^* \in X^*$. Prove that x^* is a norm-attaining complex-linear functional if and only if its real part is a norm-attaining real-linear functional.

1.114 Characterize the elements of c_0^* that are norm-attaining. Conclude that the norm-attaining functionals form a dense subset of c_0^*. (This conclusion is a special case of the *Bishop-Phelps subreflexivity theorem* of Section 2.11.)

1.115 (a) Either prove that every element of ℓ_1^* is a norm-attaining functional or display one that is not.

(b) Repeat (a) with ℓ_1^* replaced by ℓ_2^*.

1.116 (a) Let the direct sum $X_1 \oplus \cdots \oplus X_n$ of normed spaces $X_1, \dots, X_n$ be given one of the norms of Exercise 1.88. Denote this norm by $\|\cdot\|_p$. Find a norm on $X_1^* \oplus \cdots \oplus X_n^*$ that makes the map $T\colon X_1^* \oplus \cdots \oplus X_n^* \to (X_1 \oplus \cdots \oplus X_n)^*$ in the proof of Theorem 1.10.13 into an isometric isomorphism.

(b) Suppose that the norm $\|\cdot\|_1$ had been used instead of $\|\cdot\|_2$ for the norm of all direct sums of normed spaces. Let X be the real Banach space $\mathbb{R}$. Show that $(X \oplus X \oplus X)^*$ is not isometrically isomorphic to $X^* \oplus X^* \oplus X^*$. (Notice that this requires more than just showing that the mapping T defined in the proof of Theorem 1.10.13 is not an isometric isomorphism.)

1.117 Let X and Y be normed spaces.

(a) Suppose that $x^* \in X^*$ and $y \in Y$. Define $T_{x^*,y}\colon X \to Y$ by the formula $T_{x^*,y}(x) = (x^*x)y$. Prove that $T_{x^*,y} \in B(X,Y)$ and that $\|T_{x^*,y}\| = \|x^*\|\,\|y\|$.

(b) Suppose that $Y \neq \{0\}$. Prove that $B(X,Y)$ has a closed subspace isometrically isomorphic to X^*.

(c) Suppose that $X \neq \{0\}$. Prove that $B(X,Y)$ has a closed subspace isometrically isomorphic to Y.

(d) Suppose that $X \neq \{0\}$ and Y is not a Banach space. Prove that $B(X,Y)$ is not a Banach space. (See the comments following Theorem 1.4.8.)

1.118 Let X and Y be normed spaces. Prove that $B(X,Y^*)$ is isometrically isomorphic to $B(Y,X^*)$.

1.119 Exchanging the roles of X and X^* in Theorem 1.9.10 results in the following statement, which will be called statement 1.9.10*:

Suppose that X is a normed space. Let A be a nonempty subset of X^* and let $\{\, c_{x^*} : x^* \in A \,\}$ be a corresponding collection of scalars. Then the following are equivalent.

(1) *There is an x_0 in X such that $x^*x_0 = c_{x^*}$ for each x^* in A.*

(2) *There is a nonnegative real number M such that*

$$|\alpha_1 c_{x_1^*} + \cdots + \alpha_n c_{x_n^*}| \leq M\,\|\alpha_1 x_1^* + \cdots + \alpha_n x_n^*\|$$

for each linear combination $\alpha_1 x_1^* + \cdots + \alpha_n x_n^*$ of elements of A.

If (2) holds, then x_0 can be chosen in (1) so that $\|x_0\| \leq M$.

Helly's theorem is of course a weakened version of statement 1.9.10*. The purpose of this exercise is to show that statement 1.9.10* is in general false, and that in fact Helly's theorem is the best result that can be obtained in this direction, even for Banach spaces, without putting additional restrictions on X. See also Exercise 1.134.

(a) Find a Banach space X, a countable subset A of X^*, and a corresponding collection of scalars satisfying (2) but not (1) in statement 1.9.10*. (It can be done with c_0.) Thus, the requirement in Helly's theorem that the collection of functionals be finite cannot in general be relaxed.

(b) Find a Banach space X, a nonempty finite subset A of X^*, and a corresponding collection of scalars satisfying (2) for some M but for which $\|x_0\| > M$ whenever x_0 satisfies (1). Thus, the ϵ in the conclusion of Helly's theorem is in general necessary.

1.120 Let X be a normed space and let B be a subset of X^*.

(a) Prove that ${}^\perp(B^\perp) \subseteq ({}^\perp B)^\perp$.

For the rest of this exercise, let $X = c_0$ and let B be the subset of X^* that corresponds to the set $\{ (\alpha_n) : (\alpha_n) \in \ell_1, \sum_n \alpha_n = 0 \}$ when c_0^* and ℓ_1 are identified in the usual way.

(b) Show that ${}^\perp(B^\perp) = B$.

(c) Show that $({}^\perp B)^\perp = X^*$. Thus, the inclusion in (a) may be proper, even when B is a closed subspace of X^*.

1.121 Let K be a compact Hausdorff space and let $T\colon \mathrm{rca}(K) \to C(K)^*$ be the isometric isomorphism of Example 1.10.6.

(a) A linear functional F on $C(K)$ is said to be *positive* if $F(f) \geq 0$ whenever f is nonnegative-real-valued. Define a relation on $C(K)^*$ by declaring that $x_1^* \preceq x_2^*$ whenever $x_2^* - x_1^*$ is positive. Show that this relation is a partial order.

(b) Define a relation on $\mathrm{rca}(K)$ by declaring that $\mu_1 \preceq \mu_2$ whenever $\mu_2 - \mu_1$ is nonnegative-real-valued. Show that this relation is a partial order.

(c) Suppose that $\mu_1, \mu_2 \in \mathrm{rca}(K)$. Prove that $\mu_1 \preceq \mu_2$ if and only if $T\mu_1 \preceq T\mu_2$.

1.11 The Second Dual and Reflexivity

1.11.1 Definition. Let X be a normed space. The *second dual* or *double dual* or *bidual* of X is the dual space $(X^*)^*$ of X^* and is denoted by X^{**}. Similarly, the *third dual* of X is $(X^{**})^*$ and is denoted by X^{***} or $X^{(3)}$. The n^{th} *dual* $X^{(n)}$ of X is then defined inductively to be $(X^{(n-1)})^*$.

1.11.2 Example. Since c_0^* is isometrically isomorphic to ℓ_1 and ℓ_1^* is isometrically isomorphic to ℓ_∞, it follows from Theorem 1.10.12 that c_0^{**} is isometrically isomorphic to ℓ_∞. Similarly, if $1 < p < \infty$ then ℓ_p^{**} is isometrically isomorphic to ℓ_p. These isometric isomorphisms are usually treated as identifications by saying that *c_0^{**} is ℓ_∞* and that *ℓ_p^{**} is ℓ_p* if $1 < p < \infty$. However, see the warning note that ends this section.

Let x_0 be an element of a normed space X and let $Q(x_0)$ be the map from X^* into the scalar field given by the formula $\big(Q(x_0)\big)(x^*) = x^*x_0$. It is easy to check that $Q(x_0)$ is a linear functional on X^*. By Theorem 1.10.9,

$$\sup\{\,\big|\big(Q(x_0)\big)(x^*)\big| : x^* \in B_{X^*}\,\} = \sup\{\,|x^*x_0| : x^* \in B_{X^*}\,\} = \|x_0\|,$$

so $Q(x_0) \in X^{**}$ and $\|Q(x_0)\| = \|x_0\|$. If $Q(x)$ is defined similarly for each x in X, then the resulting mapping $Q\colon X \to X^{**}$ is clearly linear and so is an isometric isomorphism. Furthermore, the subspace $Q(X)$ of the Banach space X^{**} is closed if and only if it is complete, which happens if and only if X is complete. All of this is summarized in the following proposition.

1.11.3 Proposition. *Let X be a normed space and let $\big(Q(x)\big)(x^*) = x^*x$ whenever $x \in X$ and $x^* \in X^*$. Then $Q(x) \in X^{**}$ whenever $x \in X$, and Q is an isometric isomorphism from X into X^{**}. Furthermore, the subspace $Q(X)$ of X^{**} is closed if and only if X is a Banach space.*

1.11.4 Definition. The map Q in the preceding proposition is called the *natural map* or *canonical embedding map* from X into X^{**}.

Recall that if S is a metric space and C is a complete metric space that includes a dense subset isometric to S, then C is called a *completion* of S. As is shown by the proof of the next theorem, the natural map from a normed space into its second dual can be used to complete an incomplete normed space to a Banach space. This theorem also says, roughly speaking, that such a completion is unique and has the same dual space as the original normed space.

1.11.5 Theorem. *Let X be a normed space. Then there is a Banach space Y and an isometric isomorphism $T\colon X \to Y$ such that $T(X)$ is dense in Y. Furthermore, the space X^* is isometrically isomorphic to Y^*. If Z is another Banach space such that there is an isometric isomorphism from X onto a dense subset of Z, then Z is isometrically isomorphic to Y.*

PROOF. Let Q be the natural map from X into X^{**} and let $Y = \overline{Q(X)}$. Since X^{**} is a Banach space, so is its closed subspace Y, and so Q is an isometric isomorphism from X onto a dense subspace of the Banach space Y.

Define $S\colon Y^* \to X^*$ by the formula $S(y^*) = y^*Q$; that is, let $\langle x, S(y^*)\rangle = \langle Qx, y^*\rangle$ whenever $x \in X$ and $y^* \in Y^*$. It is clear that S is linear. If $x^* \in X^*$, then the bounded linear functional x^*Q^{-1} on $Q(X)$ has a bounded linear extension y^* to Y, and $Sy^* = x^*$. Thus, the operator S maps Y^* onto X^*. If $y^* \in Y^*$, then it follows from the continuity of y^* and the density of $Q(B_X)$ in B_Y that

$$\|y^*\| = \sup\{\, |\langle Qx, y^*\rangle| : x \in B_X \,\} = \sup\{\, |\langle x, Sy^*\rangle| : x \in B_X \,\} = \|Sy^*\|,$$

so S is an isometric isomorphism from Y^* onto X^*.

Finally, suppose that R is an isometric isomorphism from X onto a dense subspace of a Banach space Z. Then RQ^{-1} is an isometric isomorphism from the dense subspace $Q(X)$ of Y into Z, and so by Theorem 1.9.1 can be extended to an isometric isomorphism R_0 from Y into Z. Since $R_0(Y)$ is a Banach space that includes the dense subspace $R(X)$ of Z, it follows that $R_0(Y) = Z$, which completes the proof. ■

There are other ways to prove that every incomplete normed space can be completed to a Banach space. See Exercise 1.123 for a proof that requires quite a bit more work than that of Theorem 1.11.5, but has the advantage that it could have been given almost as soon as Banach spaces and isometric isomorphisms had been defined.

Suppose that X is a finite-dimensional normed space. It is a standard fact from linear algebra that the space of all linear functionals on X has the same finite dimension as X itself. Since each of these linear functionals is bounded, the spaces X and X^* have the same dimension, so the dimension of $X^{(n)}$ equals that of X for each positive integer n. Since the natural map Q from X into X^{**} is one-to-one, its range has the same dimension as X, which implies that Q actually maps X *onto* X^{**}.

1.11.6 Definition. (H. Hahn, 1927 [99]). A normed space X is *reflexive* if the natural map from X into X^{**} is onto X^{**}.

Actually, Hahn called such spaces *regular*. The more descriptive term *reflexive* was coined by Edgar R. Lorch [158] in 1939.

1.11.7 Theorem. *Every reflexive normed space is a Banach space.*

PROOF. Every reflexive normed space is isomorphic to a Banach space, namely, its own second dual, and so is itself a Banach space. ■

Some authors call an incomplete normed space reflexive if its completion is reflexive in the sense of Definition 1.11.6. The term *prereflexive* is also used for such a space.

1.11.8 Proposition. *Every normed space isomorphic to a reflexive normed space is itself reflexive.*

PROOF. Let T be an isomorphism from a reflexive normed space X onto a normed space Y, and let $T^*(y^*) = y^*T$ and $T^{**}(x^{**}) = x^{**}T^*$ for each y^*

in Y^* and each x^{**} in X^{**}. By Theorem 1.10.12, the maps T^* and T^{**} are isomorphisms from Y^* and X^{**} onto X^* and Y^{**} respectively. Let Q_X and Q_Y be the natural maps from X and Y into X^{**} and Y^{**} respectively. Fix an element y^{**} of Y^{**} and let x be the element of X for which $T^{**}Q_X x = y^{**}$. If $y^* \in Y^*$, then

$$\langle y^*, y^{**}\rangle = \langle y^*, T^{**}Q_X x\rangle = \langle T^*y^*, Q_X x\rangle = \langle x, T^*y^*\rangle = \langle Tx, y^*\rangle,$$

so $y^{**} = Q_Y(Tx)$. Thus, the map Q_Y is onto Y^{**}. ■

Actually, a Banach space is reflexive whenever it is the image of a reflexive normed space under any bounded linear map whatever, whether or not the map is an isomorphism. That fact appears below as Corollary 1.11.22, but Proposition 1.11.8 is used in its proof.

The discussion preceding the definition of reflexivity proves the following result.

1.11.9 Theorem. *Every finite-dimensional normed space is reflexive.*

Let n be a positive integer. As an application of Theorem 1.11.9, here is a proof that $(\ell_\infty^n)^*$ can be identified with ℓ_1^n as claimed in Example 1.10.5. Let ℓ_∞^n be identified with $(\ell_1^n)^*$ as in that example, with $T\colon \ell_\infty^n \to (\ell_1^n)^*$ being the identifying isometric isomorphism. Then

$$\big\langle (\alpha_1,\dots,\alpha_n), T(\beta_1,\dots,\beta_n)\big\rangle = \sum_{j=1}^n \alpha_j\beta_j$$

whenever $(\alpha_1,\dots,\alpha_n) \in \ell_1^n$ and $(\beta_1,\dots,\beta_n) \in \ell_\infty^n$. By Theorem 1.10.12, the map T^* given by the formula $T^*(x^{**}) = x^{**}T$ is an isometric isomorphism from $(\ell_1^n)^{**}$ onto $(\ell_\infty^n)^*$. Since ℓ_1^n is reflexive, the natural map Q from ℓ_1^n into $(\ell_1^n)^{**}$ is onto $(\ell_1^n)^{**}$, so T^*Q is an isometric isomorphism from ℓ_1^n onto $(\ell_\infty^n)^*$. Whenever $(\alpha_1,\dots,\alpha_n) \in \ell_1^n$ and $(\beta_1,\dots,\beta_n) \in \ell_\infty^n$,

$$\begin{aligned}\big\langle (\beta_1,\dots,\beta_n), T^*Q(\alpha_1,\dots,\alpha_n)\big\rangle &= \big\langle T(\beta_1,\dots,\beta_n), Q(\alpha_1,\dots,\alpha_n)\big\rangle \\ &= \big\langle (\alpha_1,\dots,\alpha_n), T(\beta_1,\dots,\beta_n)\big\rangle \\ &= \sum_{j=1}^n \alpha_j\beta_j,\end{aligned}$$

so T^*Q identifies ℓ_1^n with $(\ell_\infty^n)^*$ in the desired way.

Notice that all the preceding argument really says is that since ℓ_∞^n "is" $(\ell_1^n)^*$ and ℓ_1^n "is" $(\ell_1^n)^{**}$, it follows that ℓ_1^n "is" $(\ell_\infty^n)^*$. The role of the isometric isomorphisms Q, T, and T^* is to make the uses of the word "is" precise.

1.11.10 Theorem. *Suppose that $1 < p < \infty$. If (Ω, Σ, μ) is a positive measure space, then $L_p(\Omega, \Sigma, \mu)$ is reflexive. In particular, the space ℓ_p is reflexive.*

PROOF. Let q be conjugate to p, and let $L_p = L_p(\Omega, \Sigma, \mu)$ and $L_q = L_q(\Omega, \Sigma, \mu)$. Let $T_q \colon L_q \to L_p^*$ and $T_p \colon L_p \to L_q^*$ be the usual isometric isomorphisms as in Example 1.10.2; see also the paragraph following that example. Let Q be the natural map from L_p into L_p^{**}. Suppose that $x^{**} \in L_p^{**}$. Then $x^{**}T_q \in L_q^*$, so there is an f in L_p such that $x^{**}T_q = T_p(f)$. If $x^* \in L_p^*$, then there is a g in L_q such that $x^* = T_q(g)$, so

$$x^{**}(x^*) = x^{**}T_q(g) = \big(T_p(f)\big)(g) = \int_\Omega gf\,d\mu = \big(T_q(g)\big)(f) = x^*(f).$$

It follows that $x^{**} = Qf$, so Q is onto L_p^{**}. ■

With (Ω, Σ, μ) as in the preceding theorem, it will be seen in Example 1.11.24 that $L_1(\Omega, \Sigma, \mu)$ is reflexive if and only if it is finite-dimensional, and similarly for $L_\infty(\Omega, \Sigma, \mu)$.

It is time to give an example of a nonreflexive Banach space. The following result is a useful tool for doing so.

1.11.11 Proposition. *Let X be a reflexive normed space. Then every member of X^* is norm-attaining.*

PROOF. Let x^* be a member of X^*. By Theorem 1.10.9, there is an x^{**} in $B_{X^{**}}$ such that $|x^{**}x^*| = \|x^*\|$. If $Q \colon X \to X^{**}$ is the natural map, then there is an x in B_X such that $Qx = x^{**}$, and so $\|x^*\| = |x^{**}x^*| = |x^*x|$. ■

1.11.12 Example. By Example 1.10.10, the dual of c_0 has a member that is not norm-attaining, so c_0 is not reflexive.

For Banach spaces, the converse of Proposition 1.11.11 is true. That is, if a Banach space X has the property that every member of X^* is norm-attaining, then X is necessarily reflexive. This result, called *James's theorem,* is proved in Section 1.13.

Just as the dual space of a subspace of a normed space can be characterized in terms of the annihilator of that subspace, it is possible to characterize the second dual of the subspace in terms of its second order annihilator.

1.11.13 Definition. Suppose that A is a subset of a normed space X. The sets $A^{\perp\perp}, A^{\perp\perp\perp}, \dots$ are defined inductively to be the respective subsets $(A^\perp)^\perp, (A^{\perp\perp})^\perp, \dots$ of the respective normed spaces $X^{**}, X^{***}, \dots$ and are called the *annihilators of the second, third, ... order* of A, respectively. The abbreviation $A^{\perp(3)}$ is sometimes used for $A^{\perp\perp\perp}$, with a similar convention for annihilators of higher order.

1.11.14 Proposition. *Let M be a subspace of a normed space X. Then there is an isometric isomorphism that identifies M^{**} with $M^{\perp\perp}$ such that if an element of M^{**} is identified with the element $m^{\perp\perp}$ of $M^{\perp\perp}$ and M^* is identified with $X^*/M^\perp$ as in Theorem 1.10.16, then the action of $m^{\perp\perp}$ on M^* is given by the formula $m^{\perp\perp}(x^* + M^\perp) = m^{\perp\perp}x^*$.*

PROOF. With X^*, $M^\perp$, and $M^{\perp\perp}$ playing the respective roles of X, M, and $M^\perp$ in Theorem 1.10.17, let S be the isometric isomorphism from $M^{\perp\perp}$ onto $(X^*/M^\perp)^*$ produced by that theorem, so that

$$\langle x^* + M^\perp, Sm^{\perp\perp}\rangle = m^{\perp\perp}x^*$$

whenever $m^{\perp\perp} \in M^{\perp\perp}$ and $x^* + M^\perp \in X^*/M^\perp$. Let T be the isometric isomorphism from $X^*/M^\perp$ onto M^* of Theorem 1.10.16, so that

$$\langle m, T(x^* + M^\perp)\rangle = x^*m$$

whenever $x^* + M^\perp \in X^*/M^\perp$ and $m \in M$, and let T^* be the resulting isometric isomorphism from M^{**} onto $(X^*/M^\perp)^*$ guaranteed by Theorem 1.10.12. Then $(T^*)^{-1}S$ is the desired isometric isomorphism from $M^{\perp\perp}$ onto M^{**}. Indeed, if $m^{\perp\perp} \in M^{\perp\perp}$ and $x^* + M^\perp \in X^*/M^\perp$, then

$$\begin{aligned}\langle T(x^* + M^\perp), (T^*)^{-1}Sm^{\perp\perp}\rangle &= \langle x^* + M^\perp, T^*(T^*)^{-1}Sm^{\perp\perp}\rangle \\ &= \langle x^* + M^\perp, Sm^{\perp\perp}\rangle \\ &= m^{\perp\perp}x^*,\end{aligned}$$

as required. ■

As is suggested by the preceding proposition and Theorem 1.10.16, there is a close relationship between the higher order duals of a subspace of a normed space and the higher order annihilators of the subspace. See Exercise 1.127.

When two normed spaces are isometrically isomorphic in some natural way, it is common practice to treat the spaces as if they were the same space with two different sets of labels for its elements, and then substitute one of the spaces and its elements' labels for the other space and its elements' labels in some expression or argument. Given this "substitution principle," Proposition 1.11.14 becomes rather obvious. When M is a subspace of a normed space X, Theorems 1.10.17 and 1.10.16 provide natural ways to identify $M^{\perp\perp}$ with $(X^*/M^\perp)^*$ and $X^*/M^\perp$ with M^*, so that $M^{\perp\perp}$ can be treated as the dual space of M^*, with the action of the elements of $M^{\perp\perp}$ on those of M^* given by the formula that shows how the elements of $M^{\perp\perp}$ act on those of $X^*/M^\perp$. The proof of Proposition 1.11.14 amounts to nothing more than combining the appropriate isometric isomorphisms in the correct order to make this argument rigorous.

This substitution principle is often used without comment in the literature to shorten an argument by suppressing portions of the argument that amount only to pushing around identification maps in a fairly obvious way. This becomes especially helpful when a host of such identifying maps would be required, as is not uncommon with certain types of arguments in Banach space theory. The best attitude to take toward an argument using this substitution principle is that it is actually just a sketch of an argument whose details can be readily filled in by the concerned reader with the help of the following observations.

1. Most of the important properties that normed spaces can have, such as completeness and reflexivity, are preserved by isometric isomorphisms.

2. There are specific results saying that when a normed space is defined by some particular expression involving other normed spaces, and some of the normed spaces in the expression are replaced by others isometrically isomorphic to those replaced, then the resulting normed space is isometrically isomorphic to the original one. See Corollary 1.8.13 for an example of such a result.

3. When an isometric isomorphism T from a normed space X onto a normed space Y is treated as a natural way to identify the two spaces, the isometric isomorphisms $T^*\colon Y^* \to X^*, T^{**}\colon X^{**} \to Y^{**}$, $T^{***}\colon Y^{***} \to X^{***}$, and so forth formed by repeated applications of Theorem 1.10.12 can be considered natural identifications of the corresponding duals, second duals, and higher-order duals of X and Y.

4. With T, X, and Y as in the preceding item, any linear operator L from Y into a normed space Z can be treated as a linear operator from X into Z by identifying L with LT. The operator LT is bounded, open, one-to-one, onto Z, an isomorphism, or an isometric isomorphism if and only if L has that same property, which further justifies the identification. Analogous remarks can be made for any operator with domain X or Y or range in X or Y by preceding or following that operator with T or T^{-1}.

This substitution principle is sometimes extended to normed spaces that are only isomorphic rather than isometrically isomorphic. In any case, care should be taken not to use the principle in situations to which it does not apply. For example, the principle does not hold for denominators of quotient spaces, since it is quite possible for a Banach space X to have two isometrically isomorphic closed subspaces Y and Z such that X/Y and X/Z are not even isomorphic. See Exercise 1.85.

Suppose that M is a subspace of a normed space X and that M^{**} is identified with $M^{\perp\perp}$ as in Proposition 1.11.14. Let Q_M and Q_X be the natural maps from M and X respectively into their second duals. Then M is reflexive if and only if $Q_M(M) = M^{\perp\perp}$. It is a simple but remarkably useful fact that this statement remains true if Q_M is replaced by Q_X and

$M^{\perp\perp}$ is given its usual meaning as a subspace of X^{**}. One of the most important theorems about reflexive spaces follows almost immediately.

1.11.15 Lemma. *Let M be a subspace of a normed space X and let Q be the natural map from X into X^{**}. Then M is reflexive if and only if $Q(M) = M^{\perp\perp}$.*

PROOF. This proof uses the substitution principle discussed above. Let M^* and M^{**} be identified with $X^*/M^\perp$ and $M^{\perp\perp}$ respectively in the usual ways. Suppose first that M is reflexive. Let $m^{\perp\perp}$ be an element of $M^{\perp\perp}$ and let m^{**} be $m^{\perp\perp}$ viewed as a member of M^{**}. Then there is an m in M such that for each x^* in X^*,

$$m^{\perp\perp}x^* = \langle x^* + M^\perp, m^{**}\rangle = \langle m, x^* + M^\perp\rangle = x^*m,$$

so $Qm = m^{\perp\perp}$. This shows that $Q(M) \supseteq M^{\perp\perp}$. The reverse inclusion is easy to check, so $Q(M) = M^{\perp\perp}$.

Now suppose conversely that $Q(M) = M^{\perp\perp}$. Let m^{**} be an element of M^{**} and let $m^{\perp\perp}$ be the corresponding element of $M^{\perp\perp}$. Then there is an m in M such that $Qm = m^{\perp\perp}$, so for each member $x^* + M^\perp$ of M^*,

$$\langle x^* + M^\perp, m^{**}\rangle = m^{\perp\perp}x^* = x^*m = \langle m, x^* + M^\perp\rangle.$$

It follows that the natural map from M into M^{**} is onto M^{**}, so M is reflexive. ■

1.11.16 Theorem. (B. J. Pettis, 1938 [181]). *Every closed subspace of a reflexive normed space is reflexive.*

PROOF. Let M be a closed subspace of a reflexive normed space X and let Q be the natural map from X onto X^{**}. It is clear that $x \in {}^\perp(M^\perp)$ if and only if $Qx \in M^{\perp\perp}$, so $Q({}^\perp(M^\perp)) = M^{\perp\perp}$. But ${}^\perp(M^\perp) = M$ by Proposition 1.10.15 (c), so an appeal to the lemma completes the proof. ■

1.11.17 Corollary. (B. J. Pettis, 1938 [181]). *If X is a Banach space, then X is reflexive if and only if X^* is reflexive.*

PROOF. Suppose that X is reflexive. Let Q_X and Q_{X^*} be the natural maps from X and X^* into X^{**} and X^{***} respectively. Let x^{***} be an element of X^{***}. If $x^{**} \in X^{**}$ and $x = Q_X^{-1}x^{**}$, then

$$\langle x^{**}, x^{***}\rangle = \langle Q_X x, x^{***}\rangle = \langle x, x^{***}Q_X\rangle = \langle x^{***}Q_X, x^{**}\rangle,$$

so $x^{***} = Q_{X^*}(x^{***}Q_X)$. Since Q_{X^*} is onto X^{***}, the space X^* is reflexive.

Conversely, suppose that X^* is reflexive. Then both X^{**} and its closed subspace $Q_X(X)$ are reflexive, so X is reflexive since it is isomorphic to $Q_X(X)$. ■

1.11.18 Corollary. *If X is a reflexive normed space and M is a closed subspace of X, then X/M is reflexive.*

PROOF. The space $M^\perp$ is reflexive since it is a closed subspace of the reflexive space X^*. Since $(X/M)^* \cong M^\perp$, the space $(X/M)^*$ is reflexive. As a Banach space with a reflexive dual, the space X/M is reflexive. ■

1.11.19 Corollary. *Reflexivity is a three-space property.*

PROOF. Let M be a closed subspace of a normed space X such that both M and X/M are reflexive. It suffices to prove that X is reflexive. Let Q be the natural map from X into X^{**} and let x^{**} be a member of X^{**}. It suffices to prove that $x^{**} \in Q(X)$. Let T be the usual isometric isomorphism from $M^\perp$ onto $(X/M)^*$. Since $x^{**}T^{-1} \in (X/M)^{**}$ and X/M is reflexive, there is an $x + M$ in X/M such that whenever $m^\perp \in M^\perp$,

$$\langle m^\perp, x^{**} \rangle = \langle Tm^\perp, x^{**}T^{-1} \rangle = \langle x + M, Tm^\perp \rangle = \langle x, m^\perp \rangle = \langle m^\perp, Qx \rangle.$$

It follows that $x^{**} - Qx \in M^{\perp\perp}$. By Lemma 1.11.15, there is an m in M such that $Qm = x^{**} - Qx$, so $Q(m + x) = x^{**}$. ■

1.11.20 Corollary. *Suppose that $X_1, \dots, X_n$ are normed spaces and that $X = X_1 \oplus \cdots \oplus X_n$. Then X is reflexive if and only if each X_j is reflexive.*

PROOF. It can clearly be assumed that $n \geq 2$, and in fact that $X = X_1 \oplus X_2$; an induction argument based on the fact that $X_1 \oplus \cdots \oplus X_n \cong (X_1 \oplus \cdots \oplus X_{n-1}) \oplus X_n$ when $n \geq 3$ then gives the general case. Let $M = \{ (x_1, 0) : x_1 \in X_1 \}$, a closed subspace of X isometrically isomorphic to X_1 by Proposition 1.8.3. The quotient space X/M is isometrically isomorphic to X_2 by Proposition 1.8.5, so it is enough to prove that X is reflexive if and only if both M and X/M are reflexive. This follows from the preceding theorem and corollaries. ■

1.11.21 Corollary. *If a Banach space X is the internal direct sum of its closed subspaces $M_1, \dots, M_n$, then X is reflexive if and only if each M_j is reflexive.*

PROOF. Just notice that $X \cong M_1 \oplus \cdots \oplus M_n$ by Proposition 1.8.10 (b), then apply Corollary 1.11.20. ■

1.11.22 Corollary. *Let X be a reflexive normed space and Y a Banach space. If there is a bounded linear operator from X onto Y, then Y is reflexive.*

PROOF. Suppose that T is a bounded linear operator from X onto Y. Then $Y \cong X/\ker(T)$, and $X/\ker(T)$ is reflexive by Corollary 1.11.18. ■

Many Banach spaces can be shown to be nonreflexive by beginning with the nonreflexive space c_0 and repeatedly applying Theorem 1.11.16 and its corollaries.

1.11.23 Example. Since ℓ_1 and ℓ_∞ are isometrically isomorphic to c_0^* and c_0^{**} respectively, neither is reflexive. Another proof that ℓ_∞ is not reflexive comes from noticing that it has c_0 as a nonreflexive closed subspace.

1.11.24 Example. Suppose that (Ω, Σ, μ) is a positive measure space and that $1 \leq p \leq \infty$. If $L_p(\Omega, \Sigma, \mu)$ is infinite-dimensional, then ℓ_p is isometrically embedded in it; see Exercise 1.129. It follows immediately from the nonreflexivity of ℓ_1 that $L_1(\Omega, \Sigma, \mu)$ is reflexive if and only if it is finite-dimensional, and similarly for $L_\infty(\Omega, \Sigma, \mu)$ because of the nonreflexivity of ℓ_∞. In particular, the spaces $L_1[0,1]$ and $L_\infty[0,1]$ are not reflexive.

1.11.25 Example. Suppose that K is a compact Hausdorff space containing an infinite number of points. Let (x_n) be a sequence of distinct points in K. For each n, let δ_n be the Borel measure on K such that $\delta_n(A) = 1$ if $x_n \in A$ and $\delta_n(A) = 0$ otherwise. Define a map $T\colon \ell_1 \to \text{rca}(K)$ by the formula $T(\alpha_n) = \sum_n \alpha_n \delta_n$. Then T can be shown to embed ℓ_1 isometrically into $\text{rca}(K)$; see Exercise 1.125. Since $\text{rca}(K)$ has a nonreflexive closed subspace, it is not reflexive. Since $C(K)^*$ is isometrically isomorphic to $\text{rca}(K)$, the space $C(K)$ is also not reflexive.

The argument just given is ultimately based on Proposition 1.11.11, since that result was used to show that c_0 is not reflexive and therefore that the space ℓ_1 isometrically isomorphic to c_0^* is not reflexive. The nonreflexivity of $C(K)$ can also be obtained directly from Proposition 1.11.11. Let (x_n) and (δ_n) be as above. Since K is compact, there is an x' in K such that each neighborhood of x' contains infinitely many terms of (x_n). It can be assumed that $x' = x_1$. Let $\mu = -2^{-1}\delta_1 + \sum_{n=2}^\infty 2^{-n}\delta_n$. It is easy to check that $\|\mu\| = 1$, from which it follows that $|\int_K f\,d\mu| \leq 1$ whenever $f \in B_{C(K)}$. It will now be shown that this last inequality is always strict. Suppose to the contrary that there is an f_0 in $B_{C(K)}$ such that $\left|\int_K f_0\,d\mu\right| = 1$. There is a scalar α such that $|\alpha| = 1$ and $\int_K \alpha f_0\,d\mu = \left|\int_K f_0\,d\mu\right|$, so it can be assumed that $\int_K f_0\,d\mu = 1$. Since $|f_0(x)| \leq 1$ for each x in K, and

$$1 = \int_K f_0\,d\mu = -2^{-1} f_0(x_1) + \sum_{n=2}^\infty 2^{-n} f_0(x_n),$$

it follows that $f_0(x_1) = -1$ and $f_0(x_n) = +1$ if $n \geq 2$. But this cannot be, since it would imply that there is a neighborhood of x_1 on which the real part of f_0 is negative and in which there are infinitely many points at which f_0 takes on the value 1. Thus, if $f \in B_{C(K)}$, it must be that $\left|\int_K f\,d\mu\right| < 1$, so μ, viewed as a member of $C(K)^*$, is not norm-attaining. The space $C(K)$ is therefore not reflexive, and so the space $\text{rca}(K)$ isomorphic to $C(K)^*$ is also not reflexive.

This section ends with a word of caution about the definition of reflexivity. Since the natural map from a normed space into its second dual is an isometric isomorphism, it is often useful to think of the space as being identified with a subspace of its second dual. Having done this, it is tempting to define a normed space X to be reflexive if "$X = X^{**}$," and in fact this definition appears in print in more than one place. There is nothing objectionable about this if it is understood that "$X = X^{**}$" means that the *natural* map, not just *some* map, is an isometric isomorphism from X onto X^{**}. Stefan Banach himself inadvertently created a research problem by being imprecise in this very situation.

The prototype of all Banach space textbooks is Banach's own *Théorie des Opérations Linéaires* [13], published in 1932. The following result appears on page 189 of that book, reproduced here in translation using modern terminology.

> **Theorem 13.** *Given a separable Banach space E such that every sequence of elements of E that is bounded in norm has a subsequence weakly convergent to an element of E, the space E is isometrically isomorphic to the space E^{**} (the dual of E^*).*

Though some of these terms have yet to be defined here, their meanings are not important at the moment. An examination of Banach's proof of this theorem shows that the isometric isomorphism he had in mind is the natural map from E into E^{**}, so the conclusion of Banach's theorem is that E must be reflexive. In a note on that theorem given on page 243 of Banach's book, he made the following statement.

> The converse of Theorem 12, p. 189, is obviously false, but it is not known if it is the same for the converse of Theorem 13, p. 189, that is, if isometric isomorphism of the separable Banach space E and the space E^{**} implies, yes or no, the existence in every bounded sequence of elements of E a subsequence weakly convergent to an element of E.

Though Banach did *not* insist that the isometric isomorphism be the natural map, that is probably what he meant. Nevertheless, those who took up Banach's implied challenge to prove or disprove the converse of his Theorem 13 first had to figure out whether the converse is supposed to contain this additional hypothesis. This led naturally to the question of whether or not it really makes any difference. That is, if there is *some* isometric isomorphism from a separable Banach space E onto E^{**}, must E be reflexive? R. C. James [110] finally settled this last question in the negative in a 1951 paper by constructing a nonreflexive separable Banach space J such that there is an isometric isomorphism from J onto J^{**}. This space was also used to settle several other long-standing open questions in Banach space theory, as will be seen in Section 4.5. Incidentally, in Section 1.13, and again in Section 2.8, it will be shown that the converse of Banach's Theorem 13 is *true* if the isometric isomorphism is required to be the natural map, and in fact that the property that every bounded sequence has a weakly convergent subsequence actually characterizes the reflexive spaces among all normed spaces. The converse of Banach's Theorem 13 is therefore *false* as Banach stated it, with J being a counterexample. It pays to be careful about the definition of reflexivity.

Exercises

1.122 Give a proof of Theorem 1.10.8 based on Proposition 1.11.3.

1.123 Let X be an incomplete normed space. Supply the details for the following proof that X can be completed to a Banach space.

(a) It is a standard fact that if S is an incomplete metric space, then there is an isometry from S onto a dense subset of some complete metric space; see, for example, [65] or [172]. Let T be an isometry from X onto a dense subset of a complete metric space Y. It is enough to give Y a vector space structure and a Banach norm such that T is a linear mapping and the metric of Y is induced by that norm.

(b) Define an addition of elements of Y and a multiplication of elements of Y by scalars as follows. Whenever $y, z \in Y$ and $\alpha \in \mathbb{F}$, let (w_n) and (x_n) be sequences in X such that $T(w_n) \to y$ and $T(x_n) \to z$, then let $y + z = \lim_n T(w_n + x_n)$ and $\alpha y = \lim_n T(\alpha w_n)$. These operations are unambiguously defined.

(c) The set Y is a vector space over $\mathbb{F}$ with the operations defined in (b).

(d) The map T is a linear operator from X into Y.

(e) Let d be the metric of Y. If $y_1, y_2, z \in Y$, then $d(y_1 + z, y_2 + z) = d(y_1, y_2)$; in particular, $d(y_1 - y_2, 0) = d(y_1, y_2)$.

(f) Define $\|\cdot\|_Y : Y \to \mathbb{R}$ by the formula $\|y\|_Y = d(y, 0)$. Then $\|\cdot\|_Y$ is a norm on Y such that the metric induced by $\|\cdot\|_Y$ is d, so $\|\cdot\|_Y$ is a Banach norm.

1.124 Use Proposition 1.11.11 to prove that $L_1[0,1]$ is not reflexive.

1.125 Prove that the map T in Example 1.11.25 embeds ℓ_1 isometrically into $\mathrm{rca}(K)$.

1.126 Criticize the following "proof" that $L_2[0,1]$ is reflexive: Since the dual of $L_2[0,1]$ is $L_2[0,1]$, it immediately follows that $(L_2[0,1])^{**} = L_2[0,1]$, so $L_2[0,1]$ is reflexive.

1.127 Let M be a subspace of a normed space X.

(a) For each positive integer n, find a normed space expressed in terms of $M^{\perp(n)}$ to which the n^{th} dual $M^{(n)}$ of M is isometrically isomorphic.

(b) Suppose that M is closed. For each positive integer n, find a normed space expressed in terms of $M^{\perp(n)}$ to which $(X/M)^{(n)}$ is isometrically isomorphic.

1.128 Suppose that M is a subspace of a Banach space X and that both M and $M^{\perp}$ are reflexive. Prove that X is reflexive.

1.129 Suppose that (Ω, Σ, μ) is a positive measure space, that $1 \le p \le \infty$, and that $L_p(\Omega, \Sigma, \mu)$ is infinite-dimensional.

(a) Prove that there is a sequence of disjoint members of Σ each having nonzero measure, and that if $p \neq \infty$, then the sets can be selected so that each has finite measure.

(b) Prove that ℓ_p is isometrically embedded in $L_p(\Omega, \Sigma, \mu)$.

1.130 Complete the discussion begun in Example 1.11.25 by examining the reflexivity or nonreflexivity of $C(K)$ and $\text{rca}(K)$ when K is a compact Hausdorff space with a finite number of points.

1.131 Let X and Y be Banach spaces. Prove that if X is reflexive, then $X \cong Y$ if and only if $X^* \cong Y^*$. Prove the corresponding statement for isometric isomorphisms. Compare this to the result of Exercise 1.111.

1.132 Let X be a Banach space. Show that the kernels of the finite-rank bounded linear operators with domain X are either all reflexive or all nonreflexive.

1.133 Prove or disprove: If X is a reflexive normed space and Y is a normed space such that there is a bounded linear operator from X onto Y, then the completion of Y must be reflexive.

1.134 Let X be a normed space and let statement 1.9.10* be as in Exercise 1.119. Prove that statement 1.9.10* is true for X if and only if X is reflexive. In particular, show that if X is not reflexive then there is a subset A of X^* and a corresponding collection of scalars satisfying (2) but not (1) in statement 1.9.10*.

1.12 Separability

Recall that a topological space is *separable* if it has a countable dense subset. In this section some of the special properties of separable normed spaces are explored. The first order of business is to determine which of the most commonly encountered normed spaces are separable. For this purpose, the following result is useful.

1.12.1 Proposition.

(a) *If A is a countable subset of a normed space, then $[A]$ is separable.*

(b) *If a metric space X has an uncountable subset B such that $d(x, y) \ge \epsilon$ for some positive ϵ and each pair of distinct elements x and y of B, then X is not separable.*

PROOF. For (a), it can be assumed that $A \neq \varnothing$. Since a dense subset of $\langle A \rangle$ is also a dense subset of $[A]$, it is enough to prove that $\langle A \rangle$ is separable. Let $\mathbb{Q}_0$ be the rationals (if $\mathbb{F} = \mathbb{R}$) or the complex numbers with rational real and imaginary parts (if $\mathbb{F} = \mathbb{C}$). Let S be the subset of $\langle A \rangle$ consisting of all linear combinations of elements of A formed using only scalar coefficients from $\mathbb{Q}_0$. It is easy to check that S is countable. Suppose that $x_1, \dots, x_m \in A$ and $\alpha_1, \dots, \alpha_m \in \mathbb{F}$. For each j such that $j = 1, \dots, m$ there is a sequence $(\alpha_{j,n})_{n=1}^{\infty}$ in $\mathbb{Q}_0$ converging to α_j, and it follows from the continuity of the vector space operations that the sequence $(\alpha_{1,n}x_1 + \cdots + \alpha_{m,n}x_m)_{n=1}^{\infty}$ in S converges to $\alpha_1 x_1 + \cdots + \alpha_m x_m$. Thus, the countable set S is dense in $\langle A \rangle$, which proves (a).

For (b), just notice that X has an uncountable collection of nonempty disjoint open subsets, namely, the open balls with radius $\epsilon/2$ and center in B. Since no countable subset of X can intersect each of these open sets, the space X is not separable. ■

1.12.2 Example. The classical *Weierstrass approximation theorem* says that the polynomials are dense in $C[0,1]$; see, for example, [19] or [201]. For each nonnegative integer n, let f_n be the member of $C[0,1]$ given by the formulas $f_0(t) = 1$ and $f_n(t) = t^n$ if $n \neq 0$. It follows from the Weierstrass approximation theorem that $[\{ f_n : n = 0, 1, 2, \dots \}] = C[0,1]$, so $C[0,1]$ is separable.

1.12.3 Example. If $1 \le p < \infty$, then $L_p[0,1]$ is separable. One countable dense subset is the collection of all functions of the form

$$f = \sum_{m=0}^{2^n - 1} r_m \mathbf{I}_{[\frac{m}{2^n}, \frac{m+1}{2^n})}$$

such that n is a nonnegative integer, each r_m is rational (if $\mathbb{F} = \mathbb{R}$) or complex with rational real and imaginary parts (if $\mathbb{F} = \mathbb{C}$), and $\mathbf{I}_{[a,b)}$ is the indicator function of the interval $[a,b)$; see, for example, [242]. Another countable dense subset is found by recalling that the continuous functions on $[0,1]$ are dense in $L_p[0,1]$; see, for example, [202]. If D is a countable subset of $C[0,1]$ dense in $C[0,1]$ with its usual norm, then D is easily seen to be dense in $C[0,1]$ with the $L_p[0,1]$ norm, and so D is dense in $L_p[0,1]$. The separability of $L_p[0,1]$ thus follows from that of $C[0,1]$.

1.12.4 Example. The space $L_\infty[0,1]$ is not separable. To see this, suppose that B is the collection of all functions with domain $[0,1]$ that are indicator functions of intervals of the form $[0,t]$, where $0 \le t \le 1$. Then B is an uncountable subset of $L_\infty[0,1]$ such that $\|f - g\|_\infty = 1$ whenever f and g are different members of B.

1.12.5 Example. For each t in $[0,1]$, let δ_t be the Borel measure on $[0,1]$ such that $\delta_t(A) = 1$ if $t \in A$ and $\delta_t(A) = 0$ otherwise. Then $\{ \delta_t : t \in [0,1] \}$ is an uncountable subset of rca$[0,1]$. Since $\|\delta_{t_1} - \delta_{t_2}\| = 2$ if $t_1 \neq t_2$, the space rca$[0,1]$ is not separable.

1.12.6 Example. The space c_0 is separable, since $c_0 = [\{ e_n : n \in \mathbb{N} \}]$.

1.12.7 Example. If $1 \le p < \infty$, then $\ell_p = [\{ e_n : n \in \mathbb{N} \}]$ and so is separable. However, the space ℓ_∞ is not separable. To see this, let B be the subset of ℓ_∞ consisting of all sequences whose terms are from the set $\{0, 1\}$. Then B is uncountable and $\|(\alpha_n) - (\beta_n)\|_\infty = 1$ if (α_n) and (β_n) are different members of B.

1.12.8 Example. Every finite-dimensional normed space is the closed linear hull of a finite subset of the space and so is separable.

Most of the results about reflexivity in the preceding section have their analogs for separability, and in fact more general results are often true.

1.12.9 Proposition. *Let X be a normed space.*

(a) *If X is separable and f is a continuous map from X into a topological space S, then the range of f is a separable subset of S; in particular, every normed space isomorphic to X is separable.*

(b) *If X is separable, then each subset of X, and in particular each subspace of X, is separable.*

(c) *If A is a separable subset of X, then $\overline{A}$, $\mathrm{co}(A)$, $\overline{\mathrm{co}}(A)$, $\langle A \rangle$, and $[A]$ are all separable.*

(d) *If X is separable, then its completion is also separable.*

(e) *If X is separable and M is a closed subspace of X, then X/M is separable.*

(f) *If X is the external or internal direct sum of the normed spaces $X_1, \dots, X_n$, then X is separable if and only if each X_j is separable.*

PROOF. If X were *any* topological space with a countable dense subset D and f were a continuous map from X into a topological space S, then it would follow that $f(X) = f(\overline{D}) \subseteq \overline{f(D)}$, and so $f(X)$ would have the countable dense subset $f(D)$. This gives (a). Similarly, part (b) is just a special case of the more general fact that each subset of a separable metric space is separable.

For (c), let D_A be a countable dense subset of a separable subset A of X. Then $[D_A]$ is separable by Proposition 1.12.1 (a). Since $A \subseteq \overline{D_A} \subseteq [D_A]$, each of the sets listed in the conclusion of (c) is a subset of $[D_A]$ and so is separable. Part (d) now follows immediately from (a) and (c).

For (e), suppose that X has a countable dense subset D and that M is a closed subspace of X. Let $D_{X/M} = \{ d + M : d \in D \}$. Since

$$\|(x + M) - (d + M)\| \le \|x - d\|$$

whenever $x + M \in X/M$ and $d + M \in D_{X/M}$, it follows that $D_{X/M}$ is a countable dense subset of X/M, proving (e).

For (f), suppose that X is the external or internal direct sum of the normed spaces $X_1, \dots, X_n$. If X is separable, then each X_j is separable since it is either a subspace of X or isomorphic to a subspace of X. Conversely, suppose that each X_j is separable. For each j, let D_j be a countable dense subset of X_j. If $X = X_1 \oplus \cdots \oplus X_n$, then it is clear from the form of the direct sum norm that the countable set $D_1 \times \cdots \times D_n$ is dense in X. Suppose that X is the internal direct sum of $X_1, \ldots, X_n$. If $x_j \in X_j$ when $j = 1, \ldots, n$, then for each j there is a sequence $(d_{j,m})_{m=1}^\infty$ in D_j such that $\lim_m d_{j,m} = x_j$, and so $\lim_m (d_{1,m} + \cdots + d_{n,m}) = x_1 + \cdots + x_n$. It follows that $D_1 + \cdots + D_n$ is a countable dense subset of X, which finishes the proof of (f). ∎

1.12.10 Corollary. *Separability is a three-space property.*

PROOF. Suppose that M is a closed subspace of a normed space X such that both M and X/M are separable. Let D_M and D' be countable subsets of M and X respectively such that D_M is dense in M and $\{ x+M : x \in D' \}$ is dense in X/M. Suppose that $x_0 \in X$ and $\epsilon > 0$. Then there is an x_1 in D' such that $d(x_0 - x_1, M) = \|(x_0 + M) - (x_1 + M)\| < \epsilon/2$, and so there is an m in M such that $\|(x_0 - x_1) - m\| < \epsilon/2$. There is an x_2 in D_M such that $\|m - x_2\| < \epsilon/2$, and so $\|x_0 - (x_1 + x_2)\| = \|(x_0 - x_1) - x_2\| < \epsilon$. It follows that the countable set $D' + D_M$ is dense in X, and so X is separable. The rest of the proof follows from parts (b) and (e) of the preceding proposition. ∎

1.12.11 Theorem. *Let X be a normed space. If X^* is separable, then X is separable.*

PROOF. Let $\{ x_n^* : n \in \mathbb{N} \}$ be a countable dense subset of X^*. For each n, let x_n be an element of B_X such that $|x_n^* x_n| \geq \frac{1}{2}\|x_n^*\|$. If $x^* \in X^*$ and $x^* \neq 0$, then there is an n such that $\|x^* - x_n^*\| < \frac{1}{4}\|x^*\|$, from which it follows that $\|x_n^*\| \geq \|x^*\| - \|x^* - x_n^*\| > \frac{3}{4}\|x^*\| > 0$ and that

$$\begin{aligned} |x^* x_n| &\geq |x_n^* x_n| - |(x^* - x_n^*)(x_n)| \\ &\geq |x_n^* x_n| - \|x^* - x_n^*\| \\ &> \tfrac{1}{2}\|x_n^*\| - \tfrac{1}{4}\|x^*\| \\ &> \tfrac{1}{2}\|x_n^*\| - \tfrac{1}{3}\|x_n^*\| \\ &> 0. \end{aligned}$$

Thus, the annihilator of $\{ x_n : n \in \mathbb{N} \}$ contains only the zero element of X^*, and so

$$[\{ x_n : n \in \mathbb{N} \}] = {}^\perp(\{ x_n : n \in \mathbb{N} \}^\perp) = {}^\perp\{0\} = X.$$

An appeal to Proposition 1.12.1 (a) establishes the separability of X. ∎

The converse of the preceding theorem is false. For example, the space ℓ_1 is separable, but its dual space is isometrically isomorphic to the nonseparable space ℓ_∞ and so is not separable. The following, at least, can be said.

1.12.12 Corollary. *Let X be a reflexive normed space. Then X is separable if and only if X^* is separable.*

PROOF. The "if" portion comes from the preceding theorem. For the "only if" portion, suppose that X is separable. Then X^{**} is separable since it is isometrically isomorphic to X, so X^* is separable. ■

Notice that the corollary gives new proofs of the nonreflexivity of the spaces $L_1[0,1]$, ℓ_1, and $C[0,1]$, and therefore of $L_\infty[0,1]$, ℓ_∞, rca$[0,1]$, and c_0.

Suppose that M is a closed subspace of ℓ_1. Proposition 1.12.9 (e) guarantees that ℓ_1/M is separable. It may be somewhat surprising that, up to isomorphism, such quotients of ℓ_1 are the *only* separable Banach spaces.

1.12.13 Lemma. (S. Banach and S. Mazur, 1933 [16]). *Let X be a separable Banach space. Then there is a bounded linear operator from ℓ_1 onto X.*

PROOF. Let $\{ x_n : n \in \mathbb{N} \}$ be a countable dense subset of B_X. If $(\alpha_n) \in \ell_1$, then $\sum_n \alpha_n x_n$ is absolutely convergent and therefore convergent, so the formula $T(\alpha_n) = \sum_n \alpha_n x_n$ defines a map T from ℓ_1 into X. It is clear that T is linear. Since $\|T(\alpha_n)\| \leq \sum_n |\alpha_n| = \|(\alpha_n)\|_1$ whenever $(\alpha_n) \in \ell_1$, the operator T is bounded.

All that remains is to show that T is onto X. Suppose that $x \in B_X$. It is enough to prove that $x \in T(\ell_1)$. Select n_1 so that $\|x - x_{n_1}\| < 1/2$. Select n_2 so that $n_2 > n_1$ and $\|2(x - x_{n_1}) - x_{n_2}\| < 1/2$, that is, so that

$$\|x - x_{n_1} - 2^{-1}x_{n_2}\| < 2^{-2}.$$

Select n_3 so that $n_3 > n_2$ and $\|2^2(x - x_{n_1} - 2^{-1}x_{n_2}) - x_{n_3}\| < 1/2$, that is, so that

$$\|x - x_{n_1} - 2^{-1}x_{n_2} - 2^{-2}x_{n_3}\| < 2^{-3}.$$

Continuing in the obvious way yields a subsequence (x_{n_j}) of (x_n) such that $x = \sum_j 2^{1-j} x_{n_j}$. Let (α_n) be the element of ℓ_1 obtained by letting $\alpha_{n_j} = 2^{1-j}$ for each j and letting $\alpha_n = 0$ whenever there is no j such that $n = n_j$. Then $T(\alpha_n) = x$, so $x \in T(\ell_1)$. ■

1.12.14 Theorem. (S. Banach and S. Mazur, 1933 [16]). *For every separable Banach space X, there is a closed subspace M_X of ℓ_1 such that $X \cong \ell_1/M_X$.*

PROOF. Let T be a bounded linear operator from ℓ_1 onto X and let $M_X =$ ker(T). Then $\ell_1/M_X \cong X$ by the first isomorphism theorem for Banach spaces. ■

The final result of this section provides a characterization of separable normed spaces among all normed spaces that is occasionally useful and will have a specific application in the proof of Proposition 3.4.7.

1.12.15 Theorem. *A normed space X is separable if and only if there is a compact subset K of X such that $X = [K]$.*

PROOF. If X has a compact subset K such that $X = [K]$, then K itself has a countable dense subset A, which implies that X is separable since $X = [K] = [A]$. Suppose conversely that X is separable. It can be assumed that $X \neq \{0\}$. There is a sequence (x_n) of nonzero members of X such that

$$X = \overline{\{x_n : n \in \mathbb{N}\}} = [\{x_n : n \in \mathbb{N}\}],$$

which implies that

$$X = [\{\|nx_n\|^{-1}x_n : n \in \mathbb{N}\} \cup \{0\}].$$

This last set whose closed linear hull is being taken consists of the terms and limit of a convergent sequence in X, and so is compact. ■

A normed space that is the closed linear hull of one of its compact subsets is said to be *compactly generated.* The preceding theorem just says that a normed space is compactly generated if and only if it is separable.

Exercises

1.135 Prove that every separable infinite-dimensional normed space has a linearly independent countable dense subset.

1.136 Prove that a normed space is separable if and only if its unit sphere is separable.

1.137 Let M be a subspace of a normed space X. Prove that if M and $M^\perp$ are both separable, then X is separable.

1.138 (a) Show that no subspace of c_0 is isomorphic to ℓ_1.

(b) Show that no quotient space of c_0 is isomorphic to ℓ_1.

1.139 Let X be the subspace of ℓ_1 consisting of the sequences in ℓ_1 that sum to 0. Is X^* separable? Explain.

1.140 Give an example of a one-to-one bounded linear operator from a nonseparable normed space onto a separable normed space.

1.141 Here is a companion result for Lemma 1.12.13. Suppose that X is a Banach space and that T is a bounded linear operator from X onto ℓ_1.

(a) Let (e_n) be the sequence of standard unit vectors in ℓ_1. Prove that there is a bounded sequence (x_n) in X such that $Tx_n = e_n$ for each n.

(b) Conclude that X has a subspace isomorphic to ℓ_1.

1.142 Let K be a compact Hausdorff space. Prove that $\text{rca}(K)$ is separable if and only if K is countable.

1.143 Suppose that (Ω, Σ, μ) is a positive measure space. Prove that $L_\infty(\Omega, \Sigma, \mu)$ is separable if and only if it is finite-dimensional.

1.144 Let X be a normed space and let Z be a separable subspace of X^*. Prove that there is a separable closed subspace Y of X such that Z is isometrically isomorphic to a subspace of Y^*.

1.145 The evidence accumulated so far might lead one to conjecture that all reflexive normed spaces are separable. Give a counterexample. (Consider counting measures on uncountable sets.)

*1.13 Characterizations of Reflexivity

The purpose of this section is to obtain some characterizations of reflexive normed spaces among all normed spaces and among all Banach spaces. The culmination of these efforts will be James's theorem, a very beautiful and enormously useful result that is one of the major theorems of Banach space theory.

Several of the proofs that are about to be given are much easier, at least notationally, if done only for real normed spaces. Fortunately, the case for complex normed spaces usually follows immediately from this next result.

1.13.1 Proposition. *Let X be a complex normed space and let X_r be the real normed space obtained from X by restricting multiplication of vectors by scalars to $\mathbb{R} \times X$. Then X is reflexive if and only if X_r is reflexive.*

PROOF. For each x^* in $(X^*)_r$, let $Tx^* = \operatorname{Re} x^*$. It is an easy consequence of Proposition 1.9.3 that T is an isometric isomorphism from $(X^*)_r$ onto $(X_r)^*$. By Theorem 1.10.12, the map $T^*\colon (X_r)^{**} \to \big((X^*)_r\big)^*$ given by the formula $T^*(u^{**}) = u^{**}T$ is an isometric isomorphism from $(X_r)^{**}$ onto $\big((X^*)_r\big)^*$. Since two linear functionals on a complex vector space are equal if and only if their real parts agree on the space, it is easy to see that each of statements 1, 2, 3, and 4 below is equivalent to the statement following it.

1. The space X is reflexive.
2. For each element x^{**} of X^{**} there is an element x of X such that $\operatorname{Re}\langle x, x^*\rangle = \operatorname{Re}\langle x^*, x^{**}\rangle$ whenever $x^* \in X^*$.
3. For each element w^* of $\big((X^*)_r\big)^*$ there is an element x of X_r such that $\langle x, Tx^*\rangle = \langle x^*, w^*\rangle$ whenever $x^* \in (X^*)_r$.
4. For each element u^{**} of $(X_r)^{**}$ there is an element x of X_r such that $\langle x, Tx^*\rangle = \langle x^*, T^*u^{**}\rangle = \langle Tx^*, u^{**}\rangle$ whenever $x^* \in (X^*)_r$.
5. The space X_r is reflexive.

The equivalence of statements 1 and 5 yields the conclusion of the proposition. ∎

The portion of this section extending from here through Example 1.13.7 is devoted primarily to obtaining several closely related characterizations of reflexivity, one of which is in terms of the following type of convergence.

1.13.2 Definition. Let x be an element and (x_n) a sequence in a normed space X. Then (x_n) *converges weakly* to x if $x^*x_n \to x^*x$ whenever $x^* \in X^*$.

It will be shown in Chapter 2 that every normed space has a topology $\mathfrak{T}$ such that a sequence in the space converges weakly to an element of the space if and only if the sequence converges to that element with respect to $\mathfrak{T}$. For the moment, the statement that a sequence converges weakly to a certain limit should not be taken to imply anything more than is stated in Definition 1.13.2. Notice that if a sequence (x_n) in a normed space X converges weakly to elements x and y of X, then $x^*x = \lim_n x^*x_n = x^*y$ for each x^* in X^*, and so $x = y$ by Corollary 1.9.9. That is, no sequence in a normed space has more than one "weak limit." Notice also that the convergence of a sequence in a normed space in the usual sense implies the weak convergence of that sequence to the same limit. Incidentally, the converse is not true. The sequence (e_n) of standard unit vectors in ℓ_2 is obviously weakly convergent to 0 and equally obviously not convergent to anything with respect to the norm topology since the sequence is not Cauchy.

The following lemma will be superseded by Theorems 1.13.5 and 1.13.6, in which it will be shown that statements (a), (b), and (c) of this lemma are actually equivalent.

1.13.3 Lemma. *Let X be a normed space. Then* (a) $\Rightarrow$ (b) $\Rightarrow$ (c) *in the following collection of statements.*

(a) *The space X is reflexive.*

(b) *Every bounded sequence in X has a weakly convergent subsequence.*

(c) *Whenever (C_n) is a sequence of nonempty closed bounded convex sets in X such that $C_n \supseteq C_{n+1}$ for each n, it follows that $\bigcap_n C_n \neq \varnothing$.*

PROOF. Suppose that (b) holds. Let (C_n) be a sequence of nonempty closed bounded convex subsets of X such that $C_1 \supseteq C_2 \supseteq \cdots$. For each positive integer n, let x_n be an element of C_n, and let x be the limit of a weakly convergent subsequence (x_{n_k}) of (x_n). If $x \notin C_m$ for some m, then by Proposition 1.9.15 there is an x^* in X^* such that

$$\operatorname{Re} x^*x > \sup\{\operatorname{Re} x^*y : y \in C_m\} \geq \lim_k \operatorname{Re} x^*x_{n_k} = \operatorname{Re} x^*x,$$

a contradiction. Therefore $x \in \bigcap_n C_n$, which shows that (b) $\Rightarrow$ (c).

Now suppose that it has been proved that (a) $\Rightarrow$ (b) when X is separable. For the general case, let X be reflexive and let (x_n) be a bounded sequence in X. Let $Y = [\{x_n : n \in \mathbb{N}\}]$, a separable reflexive subspace of X. Then there is a subsequence (x_{n_k}) of (x_n) and an x in Y such that $y^*x_{n_k} \to y^*x$ whenever $y^* \in Y^*$. Since the restriction of each member of X^* to Y is in Y^*, it follows that (x_{n_k}) converges weakly to x in X. Therefore all that is left to be proved is that (a) $\Rightarrow$ (b) when X is separable.

Let X be a separable reflexive normed space and let $\{ y_n^* : n \in \mathbb{N} \}$ be a countable dense subset of X^*. Suppose that (x_n) is a bounded sequence in X. By diagonalization, a subsequence (x_{n_k}) of (x_n) can be found such that $\lim_k y_j^* x_{n_k}$ exists for each j. If $x^* \in X^*$ and $j, k, l \in \mathbb{N}$, then

$$\begin{aligned} |x^*x_{n_k} - x^*x_{n_l}| &\le |y_j^*x_{n_k} - y_j^*x_{n_l}| + |(x^* - y_j^*)(x_{n_k} - x_{n_l})| \\ &\le |y_j^*x_{n_k} - y_j^*x_{n_l}| + \|x^* - y_j^*\| \sup\{\|x_n - x_m\| : n, m \in \mathbb{N}\}. \end{aligned}$$

This and the density of $\{ y_n^* : n \in \mathbb{N} \}$ in X^* together imply that the sequence $(x^*x_{n_k})$ is Cauchy and hence convergent whenever $x^* \in X^*$. The map $x^* \mapsto \lim_k x^*x_{n_k}$ is clearly a linear functional on X^*, and this functional is in X^{**} since $|\lim_k x^*x_{n_k}| \le \|x^*\| \sup\{ \|x_n\| : n \in \mathbb{N} \}$ for each x^* in X^*. The reflexivity of X yields an x in X such that $\lim_k x^*x_{n_k} = x^*x$ whenever $x^* \in X^*$, so (x_{n_k}) is a weakly convergent subsequence of (x_n). Therefore (a) $\Rightarrow$ (b) when X is separable, which finishes the proof. ■

Most of the theorems of this section are ultimately derived from the following result.

1.13.4 Theorem. (R. C. James, 1964 [112, 114]). *Let X be a Banach space. Then the following are equivalent.*

(a) *The space X is not reflexive.*

(b) *For each θ such that $0 < \theta < 1$ there is a sequence (x_n^*) in S_{X^*} and a sequence (x_n) in S_X such that $\operatorname{Re} x_n^* x_j \ge \theta$ if $n \le j$ and $\operatorname{Re} x_n^* x_j = 0$ if $n > j$.*

(c) *For some θ such that $0 < \theta < 1$ there is a sequence (x_n^*) in S_{X^*} and a sequence (x_n) in S_X as in* (b).

PROOF. It may be assumed that X is a real Banach space, since the case for complex scalars follows from that for real scalars by Proposition 1.13.1. Suppose that (c) holds. For each positive integer n, let

$$C_n = \overline{\mathrm{co}}(\{x_n, x_{n+1}, x_{n+2}, \dots\}).$$

Notice that each C_n lies in B_X and so is bounded. Suppose that $x \in C_{n_0}$ for some n_0. If $\epsilon > 0$, then there is a nonnegative integer m_ϵ and a y in $\mathrm{co}(\{x_{n_0}, \dots, x_{n_0+m_\epsilon}\})$ such that $\|x - y\| < \epsilon$, so $|x_n^*x| = |x_n^*(x - y)| < \epsilon$ whenever $n > n_0 + m_\epsilon$. Therefore $\lim_n x_n^*x = 0$. Since $x_{n_0}^*x_j \ge \theta$ when $j \ge n_0$, it follows that $x_{n_0}^*z \ge \theta$ whenever $z \in \mathrm{co}(\{x_{n_0}, x_{n_0+1}, \dots\})$, so $x_{n_0}^*x \ge \theta$. In particular, if $x \in \bigcap_n C_n$, then $x_n^*x \ge \theta$ for each n even though

$\lim_n x_n^* x = 0$, a contradiction that shows that $\bigcap_n C_n = \varnothing$. It follows from the preceding lemma that X cannot be reflexive, so (c) $\Rightarrow$ (a).

James attributes the following proof that (a) $\Rightarrow$ (b) to Mahlon Day. Suppose that (a) holds, that is, that X is not reflexive. Fix a θ in $(0,1)$. Let Q be the natural map from X into X^{**}. Since $Q(X)$ is a closed proper subspace of X^{**}, there is a member $x^{**} + Q(X)$ of $X^{**}/Q(X)$ such that

$$\theta < \|x^{**} + Q(X)\| = d\big(x^{**}, Q(X)\big) < 1.$$

By Proposition 1.7.6, it may be assumed that $\theta < \|x^{**}\| < 1$. It will now be shown that there are sequences (x_n^*) in B_{X^*} and (x_n) in B_X such that $x_n^* x_j = \theta$ if $n \le j$, $x_n^* x_j = 0$ if $n > j$, and $x^{**} x_n^* = \theta$ for every n, from which it immediately follows that (b) holds. Let x_1^* in B_{X^*} be such that $x^{**}x_1^* = \theta$. Since $\theta \le \|x^{**}\|\,\|x_1^*\| < \|x_1^*\|$, there is an x_1 in B_X such that $x_1^* x_1 = \theta$. The rest of the sequence is constructed inductively. Suppose that $x_1, \dots, x_{n-1}$ and $x_1^*, \dots, x_{n-1}^*$ have been chosen to satisfy the required conditions. Let $M = \theta / d\big(x^{**}, Q(X)\big)$. Notice that $0 < M < 1$. If $\alpha_1, \dots, \alpha_{n-1}$ are scalars, then

$$M \left\| x^{**} + \sum_{j=1}^{n-1} \alpha_j Q x_j \right\| \ge M \cdot d\big(x^{**}, Q(X)\big) = \theta.$$

Let $c_n = \theta$ and let $c_1 = \cdots = c_{n-1} = 0$. Then for each linear combination $\alpha_1 Q x_1 + \cdots + \alpha_{n-1} Q x_{n-1} + \alpha_n x^{**}$ of $Qx_1, \dots, Qx_{n-1}, x^{**}$,

$$|\alpha_1 c_1 + \cdots + \alpha_n c_n| = |\alpha_n|\theta \le M\, \|\alpha_1 Q x_1 + \cdots + \alpha_{n-1} Q x_{n-1} + \alpha_n x^{**}\|.$$

By Helly's theorem, there is for each positive ϵ a y_ϵ^* in X^* such that

(1) $\|y_\epsilon^*\| \le M + \epsilon$;
(2) $(Qx_j)(y_\epsilon^*) = c_j = 0$ when $j = 1, \dots, n-1$; and
(3) $x^{**} y_\epsilon^* = c_n = \theta$.

Let $x_n^* = y_\epsilon^*$ for a positive ϵ small enough that $\|x_n^*\| \le 1$. Then $x^{**}x_n^* = \theta$ and $x_n^* x_j = 0$ when $j = 1, \dots, n-1$, so to finish the induction all that is needed is an x_n in B_X such that $x_j^* x_n = \theta$ when $j = 1, \dots, n$. Let $\alpha_1, \dots, \alpha_n$ be scalars. Then

$$\left| \sum_{j=1}^{n} \alpha_j \theta \right| = \left| \sum_{j=1}^{n} \alpha_j x^{**} x_j^* \right| \le \|x^{**}\| \left\| \sum_{j=1}^{n} \alpha_j x_j^* \right\|,$$

so another application of Helly's theorem produces an appropriate x_n. This completes the induction step and shows that (a) $\Rightarrow$ (b). It is obvious that (b) $\Rightarrow$ (c), so the proof is finished. ■

Theorem 1.13.4 is sometimes called *James's sequential characterization of reflexivity*. A number of other useful characterizations of reflexivity follow almost immediately from it, as will now be seen.

It is a corollary of Theorem 1.13.4 and its proof that (c) $\Rightarrow$ (a) in Lemma 1.13.3. To see this, suppose that X is a normed space that satisfies part (c) of that lemma; that is, for which $\bigcap_n C_n \neq \varnothing$ whenever (C_n) is a sequence of nonempty closed bounded convex subsets of X such that $C_n \supseteq C_{n+1}$ for each n. The goal is to prove that X is reflexive, so by Proposition 1.13.1 it may be assumed that $\mathbb{F} = \mathbb{R}$ as in the proof of Theorem 1.13.4. Let (y_n) be a Cauchy sequence in X and let $D_n = \overline{\text{co}}\{ y_j : j \geq n \}$ for each n. Then for each positive ϵ there is a positive integer n_ϵ such that D_n lies in a closed ball of radius ϵ when $n \geq n_\epsilon$. It follows that $\bigcap_n D_n$ has exactly one element y_0 and that $y_n \to y_0$. The space X is therefore a Banach space. If X were not reflexive, then part (c) of Theorem 1.13.4 would hold, and a peek at the first paragraph of the proof of that theorem shows that X would have a sequence (C_n) of nonempty closed bounded convex subsets such that $C_n \supseteq C_{n+1}$ for each n and $\bigcap_n C_n = \varnothing$. This contradiction proves that X is reflexive.

Since (c) $\Rightarrow$ (a) in Lemma 1.13.3, statements (a), (b), and (c) in that lemma are actually equivalent. This is the content of the next two theorems.

1.13.5 Theorem. *A normed space is reflexive if and only if each of its bounded sequences has a weakly convergent subsequence.*

1.13.6 Theorem. (V. L. Šmulian, 1939 [223]). *A normed space X is reflexive if and only if $\bigcap_n C_n \neq \varnothing$ whenever (C_n) is a sequence of nonempty closed bounded convex subsets of X such that $C_n \supseteq C_{n+1}$ for each n.*

1.13.7 Example. Let $\mathbb{D}$ be the closed unit disc in the complex plane and let $A(\mathbb{D})$ be the *disc algebra*, that is, the subspace of $C(\mathbb{D})$ consisting of the members of $C(\mathbb{D})$ analytic in the open unit disc. Since the uniform limit of a sequence of functions analytic in the open unit disc is analytic in that disc, it follows that $A(\mathbb{D})$ is a closed subspace of $C(\mathbb{D})$ and so is a Banach space.

For each positive integer n, define x_n in $A(\mathbb{D})$ by the formula $x_n(z) = z^n$, and let $C_n = \overline{\text{co}}(\{x_n, x_{n+1}, \dots\})$. If x is in $\text{co}(\{x_1, x_2, \dots\})$, then $\|x\|_\infty = 1$ since $x(1) = 1$ and $|x(z)| \leq 1$ for each z in $\mathbb{D}$. It follows that $C_n \subseteq S_{A(\mathbb{D})}$ for each n.

If $n \in \mathbb{N}$ and $x \in \text{co}(\{x_n, x_{n+1}, \dots\})$, then x and its first $n-1$ derivatives all have value 0 when $z = 0$. Now if (y_j) is a sequence in $A(\mathbb{D})$ and y is an element of $A(\mathbb{D})$ such that $\|y_j - y\|_\infty \to 0$, then a standard result about the uniform convergence of analytic functions assures that $y_j'(0) \to y'(0)$, $y_j''(0) \to y''(0)$, $y_j'''(0) \to y'''(0)$, and so forth. It follows that if $n \in \mathbb{N}$ and $x \in C_n$, then x and its first $n-1$ derivatives all have value 0 when $z = 0$.

Suppose that $x \in \bigcap_n C_n$. Then x and its derivatives of all orders have value 0 when $z = 0$, so a moment's thought about the power series expansion for x about 0 shows that $x = 0$, even though $\|x\|_\infty = 1$. This contradiction shows that $\bigcap_n C_n = \varnothing$, so $A(\mathbb{D})$ is not reflexive.

By Theorem 1.11.16, if a normed space is reflexive then it passes that property on to each of its closed subspaces. Conversely, if each closed subspace of a normed space X is reflexive, then X is obviously reflexive as a closed subspace of itself. It turns out, however, that it is only necessary to check all of the *separable* closed subspaces of X for reflexivity in order to conclude that X is reflexive. That is, reflexivity is a property for normed spaces that is *separably determined.*

1.13.8 Theorem. *A normed space is reflexive if and only if each of its separable closed subspaces is reflexive.*

PROOF. Suppose that the normed space X is not reflexive. It is enough to find a separable closed subspace of X that is not reflexive. Let (x_n) be a bounded sequence in X with no weakly convergent subsequence, and let $Y = [\{ x_n : n \in \mathbb{N} \}]$. Since the restriction of each member of X^* to Y is in Y^*, a subsequence of (x_n) converging weakly in Y would also converge weakly in X, so (x_n) has no subsequence converging weakly in Y. Therefore Y is the desired nonreflexive separable closed subspace of X. ∎

Theorem 1.13.6 is especially interesting because it gives a purely intrinsic test for reflexivity; that is, it requires knowledge only of the behavior of objects in X and no knowledge at all of X^* or X^{**}. This next result is also of that type.

1.13.9 Theorem. (R. C. James, 1964 [114]). *Let X be a Banach space. Then the following are equivalent.*

(a) *The space X is not reflexive.*

(b) *For each θ such that $0 < \theta < 1$ there is a sequence (x_n) in S_X such that $d\big(\mathrm{co}(\{x_1, \dots, x_n\}), \mathrm{co}(\{x_{n+1}, x_{n+2}, \dots\})\big) \geq \theta$ for each n.*

(c) *For some θ such that $0 < \theta < 1$ there is a sequence (x_n) in S_X as in* (b).

PROOF. Suppose that (a) holds. Fix a θ in $(0,1)$ and let (x_n^*) and (x_n) be as in Theorem 1.13.4 (b). If $n \in \mathbb{N}$, $y \in \mathrm{co}(\{x_1, \dots, x_n\})$, and $z \in \mathrm{co}(\{x_{n+1}, x_{n+2}, \dots\})$, then $\|y - z\| \geq \mathrm{Re}\, x_{n+1}^*(z - y) \geq \theta$, and so

$$d\big(\mathrm{co}(\{x_1, \dots, x_n\}), \mathrm{co}(\{x_{n+1}, x_{n+2}, \dots\})\big) \geq \theta.$$

This shows that (a) $\Rightarrow$ (b). It is clear that (b) $\Rightarrow$ (c).

Finally, suppose that θ and (x_n) are as in (c). For each positive integer n, let $C_n = \overline{\mathrm{co}}(\{x_{n+1}, x_{n+2}, \dots\})$, a subset of B_X. Suppose that $x \in \bigcap_n C_n$. Then there is an m in $\mathbb{N}$ and a y in $\mathrm{co}(\{x_1, \dots, x_m\})$ such that $\|x - y\| < \theta/2$, as well as a z in $\mathrm{co}(\{x_{m+1}, x_{m+2}, \dots\})$ such that $\|x - z\| < \theta/2$, which implies that $\theta \leq \|y - z\| \leq \|y - x\| + \|x - z\| < \theta$. This contradiction shows that $\bigcap_n C_n = \varnothing$, so X is not reflexive. This proves that (c) $\Rightarrow$ (a). ∎

The preceding theorem could have been stated with the convex hulls replaced by closed convex hulls, since the distance between two nonempty subsets of a metric space is the same as the distance between their closures.

Suppose that X is a Banach space. By Proposition 1.11.11, if X is reflexive then each member of X^* is norm-attaining. The converse is also true and is known as *James's theorem.* This result has a long and interesting history, most of it associated with the name of Robert C. James.

Stanisław Mazur [163] was the first to ask if a Banach space must be reflexive when all members of its dual space are norm-attaining. Though Mazur's paper appeared in 1933, the first substantial progress toward answering his question was not made until 1950, when James [108] showed that if a separable Banach space X has a *Schauder basis,* a certain type of sequence that will be discussed in Chapter 4, then X is reflexive if each Banach space Y isomorphic to X has the property that each element of Y^* is norm-attaining. In the same year Victor Klee [132] used an argument based on Theorem 1.13.6 to improve James's result by removing the requirements that X have a Schauder basis and be separable. Incidentally, Klee's paper contains a number of interesting characterizations of reflexivity that will not be covered here. Most of the paper is accessible after reading Chapters 1 and 2 of this book.

In 1957, James [111] showed that a separable Banach space is reflexive if each member of its dual space is norm-attaining. Though the same result for nonseparable Banach spaces would not appear for another seven years, the impact of James's 1957 result on Banach space theory was immediate and substantial. It is often possible to prove from this result that a Banach space is reflexive by showing that all of its separable closed subspaces are reflexive. For example, it can be shown in this way that a Banach space is reflexive whenever each of its nonempty closed convex subsets has a point nearest the origin; see Exercise 1.153. It is not at all a coincidence that the branch of approximation theory called *Banach space nearest point theory,* dealing with points of sets in Banach spaces nearest other sets or points, began to grow rapidly at about this time. The interested reader might want to look at Section 4 of the 1958 paper by Ky Fan and Irving Glicksberg [77] on spheres in normed spaces for another good example of an argument in which James's 1957 result is used to prove that a possibly nonseparable Banach space is reflexive.

James [112] finally completed his quest by showing in a 1964 paper that the separability hypothesis in his 1957 result is unnecessary. In fact, in another 1964 paper James [115] proved a stronger result called *James's weak compactness theorem,* or often just *James's theorem* since the reflexivity theorem that also goes by that name turns out to be a special case of the weak compactness theorem. It would not be easy to state James's weak compactness theorem in its full generality here, since the language needed to do so is not developed until Chapter 2. However, it is already possible to state the following version of it. Though the class of sets in the hypotheses is smaller than in the general version, this version is still strong enough to imply the reflexivity theorem. See Exercise 1.154.

Let C be a closed convex subset of a Banach space X. Suppose that whenever $x^* \in X^*$, the supremum of $|x^*|$ on C is actually attained by $|x^*|$ somewhere on C. Then every sequence in C has a weakly convergent subsequence.

The general form of James's weak compactness theorem will be derived in Section 2.9.

James's 1964 proof of his reflexivity theorem is based on Theorem 1.13.4 and is somewhat intricate. The proof given here is a greatly simplified one, also by James [117], that appeared in 1972. The plan of attack is to prove James's theorem for separable Banach spaces, then obtain the general case from that. The following technical lemma is needed for the separable case. This lemma is stated and proved in a bit more generality than is actually needed in this section, for the only immediate application will be to the case in which the set A mentioned in the lemma is the closed unit ball of the space. However, the more general result is no more difficult to prove, and will have an important application in Section 2.9.

1.13.10 Lemma. (R. C. James, 1972 [117]). *Let A be a nonempty subset of the closed unit ball of a normed space X. Suppose that (β_n) is a sequence of positive numbers with sum 1, that $0 < \theta < 1$, and that (x_n^*) is a sequence in B_{X^*} such that $\sup\{\,|x^*x| : x \in A\,\} \geq \theta$ whenever $x^* \in \mathrm{co}(\{\,x_n^* : n \in \mathbb{N}\,\})$. Then there is an α such that $\theta \leq \alpha \leq 1$ and a sequence (y_n^*) in B_{X^*} such that*

(a) *$y_n^* \in \mathrm{co}(\{\,x_j^* : j \geq n\,\})$ for each positive integer n;*

(b) *$\sup\{\,|\sum_{j=1}^{\infty} \beta_j y_j^* x| : x \in A\,\} = \alpha$; and*

(c) *$\sup\{\,|\sum_{j=1}^{n} \beta_j y_j^* x| : x \in A\,\} < \alpha(1 - \theta\sum_{j=n+1}^{\infty} \beta_j)$ for each positive integer n.*

PROOF. This proof consists of the construction by induction of the sequence (y_n^*) and a sequence (α_n) of scalars converging to α, followed by the verification of some good-sized inequalities. Here are the steps.

1. For notational convenience, let $|x^*|_A = \sup\{\,|x^*x| : x \in A\,\}$ whenever $x^* \in X^*$. It is easy to check that $|\cdot|_A$ is a continuous seminorm on X^* such that $|x^*|_A \leq \|x^*\|$ for each x^* in X^*. Let (ϵ_n) be a sequence of positive reals converging to 0 such that

$$\sum_{k=1}^{\infty} \frac{\beta_k \epsilon_k}{\sum_{j=k+1}^{\infty} \beta_j \sum_{j=k}^{\infty} \beta_j} < 1 - \theta.$$

A sequence (y_n^*) in X^* will be constructed inductively such that for each positive integer n,

$$y_n^* \in \mathrm{co}(\{\,x_j^* : j \geq n\,\}) \subseteq B_{X^*}.$$

and

$$\left|\sum_{j=1}^{n-1} \beta_j y_j^* + \left(\sum_{j=n}^{\infty} \beta_j\right) y_n^*\right|_A < \alpha_n(1+\epsilon_n), \tag{1.4}$$

where

$$\alpha_n = \inf\left\{ \left|\sum_{j=1}^{n-1} \beta_j y_j^* + \left(\sum_{j=n}^{\infty} \beta_j\right) y^*\right|_A : y^* \in \operatorname{co}(\{ x_j^* : j \ge n \}) \right\}. \tag{1.5}$$

The sums to $n-1$ in (1.4) and (1.5) are to be considered to be the zero element of X^* when $n = 1$.

2. To start the induction, notice that

$$\alpha_1 = \inf\{ |y^*|_A : y^* \in \operatorname{co}(\{ x_j^* : j \ge 1 \}) \} \ge \theta > 0,$$

so there is a y_1^* in $\operatorname{co}(\{ x_j^* : j \ge 1 \})$ such that

$$|y_1^*|_A < \alpha_1(1+\epsilon_1),$$

that is, such that (1.4) is satisfied when $n = 1$.

3. Suppose that $n \ge 2$ and that $y_1^*, \ldots, y_{n-1}^*$ have been found. If $y^* \in \operatorname{co}(\{ x_j^* : j \ge n \})$, then

$$\begin{aligned} &\sum_{j=1}^{n-1} \beta_j y_j^* + \left(\sum_{j=n}^{\infty} \beta_j\right) y^* \\ &\quad = \sum_{j=1}^{n-2} \beta_j y_j^* + \left(\sum_{j=n-1}^{\infty} \beta_j\right)\left(\frac{\beta_{n-1}}{\sum_{j=n-1}^{\infty} \beta_j} y_{n-1}^* + \frac{\sum_{j=n}^{\infty} \beta_j}{\sum_{j=n-1}^{\infty} \beta_j} y^*\right), \end{aligned}$$

where the sum to $n-2$ is to be considered to be the zero element of X^* if $n = 2$. The object inside the rightmost pair of parentheses is a convex combination of two members of $\operatorname{co}(\{ x_j^* : j \ge n-1 \})$ and so lies in $\operatorname{co}(\{ x_j^* : j \ge n-1 \})$. It follows that the set whose infimum determines α_n is a subset of the one whose infimum determines α_{n-1}, so $\alpha_{n-1} \le \alpha_n$. Therefore $\alpha_n > 0$, and so a y_n^* can be found in $\operatorname{co}(\{ x_j^* : j \ge n \})$ satisfying (1.4). This finishes the induction.

4. For each positive integer n, the set whose infimum determines α_n is easily seen to be bounded from above by 1. Since

$$\theta \le \alpha_1 \le \alpha_2 \le \cdots \le 1,$$

the sequence (α_n) converges to some α such that $\theta \le \alpha \le 1$.

5. For each positive integer n,

$$\alpha_n \leq \left| \sum_{j=1}^{n-1} \beta_j y_j^* + \left(\sum_{j=n}^{\infty} \beta_j \right) y_n^* \right|_A < \alpha_n(1+\epsilon_n),$$

so letting n tend to infinity shows that $\alpha = \left|\sum_{j=1}^{\infty} \beta_j y_j^*\right|_A$.

6. All that is left is to show that α and (y_n^*) satisfy part (c) of the conclusion of the lemma. Fix a positive integer n. If $n \geq 2$, then

$$\begin{aligned}
&\left| \sum_{j=1}^{n} \beta_j y_j^* \right|_A \\
&\quad = \left| \frac{\beta_n}{\sum_{j=n}^{\infty} \beta_j} \left(\sum_{j=1}^{n-1} \beta_j y_j^* + \left(\sum_{j=n}^{\infty} \beta_j \right) y_n^* \right) + \frac{\sum_{j=n+1}^{\infty} \beta_j}{\sum_{j=n}^{\infty} \beta_j} \sum_{j=1}^{n-1} \beta_j y_j^* \right|_A \\
&\quad \leq \frac{\beta_n}{\sum_{j=n}^{\infty} \beta_j} \left| \sum_{j=1}^{n-1} \beta_j y_j^* + \left(\sum_{j=n}^{\infty} \beta_j \right) y_n^* \right|_A + \frac{\sum_{j=n+1}^{\infty} \beta_j}{\sum_{j=n}^{\infty} \beta_j} \left| \sum_{j=1}^{n-1} \beta_j y_j^* \right|_A \\
&\quad < \frac{\beta_n}{\sum_{j=n}^{\infty} \beta_j} \alpha_n (1+\epsilon_n) + \frac{\sum_{j=n+1}^{\infty} \beta_j}{\sum_{j=n}^{\infty} \beta_j} \left| \sum_{j=1}^{n-1} \beta_j y_j^* \right|_A \\
&\quad = \left(\sum_{j=n+1}^{\infty} \beta_j \right) \left(\frac{\beta_n \alpha_n (1+\epsilon_n)}{\sum_{j=n+1}^{\infty} \beta_j \sum_{j=n}^{\infty} \beta_j} + \frac{\left|\sum_{j=1}^{n-1} \beta_j y_j^*\right|_A}{\sum_{j=n}^{\infty} \beta_j} \right),
\end{aligned}$$

giving an upper bound for $\left|\sum_{j=1}^{n} \beta_j y_j^*\right|_A$ in terms of $\left|\sum_{j=1}^{n-1} \beta_j y_j^*\right|_A$. If $n \geq 3$, then $\left|\sum_{j=1}^{n-1} \beta_j y_j^*\right|_A$ is bounded from above by an analogous expression involving $\left|\sum_{j=1}^{n-2} \beta_j y_j^*\right|_A$, and so forth. Therefore if $n \geq 2$, then

$$\begin{aligned}
\left| \sum_{j=1}^{n} \beta_j y_j^* \right|_A &< \left(\sum_{j=n+1}^{\infty} \beta_j \right) \left(\frac{\beta_n \alpha_n (1+\epsilon_n)}{\sum_{j=n+1}^{\infty} \beta_j \sum_{j=n}^{\infty} \beta_j} \right. \\
&\qquad \left. + \frac{\beta_{n-1} \alpha_{n-1} (1+\epsilon_{n-1})}{\sum_{j=n}^{\infty} \beta_j \sum_{j=n-1}^{\infty} \beta_j} + \frac{\left|\sum_{j=1}^{n-2} \beta_j y_j^*\right|_A}{\sum_{j=n-1}^{\infty} \beta_j} \right) \\
&< \cdots \\
&< \left(\sum_{j=n+1}^{\infty} \beta_j \right) \left(\sum_{k=2}^{n} \left\{ \frac{\beta_k \alpha_k (1+\epsilon_k)}{\sum_{j=k+1}^{\infty} \beta_j \sum_{j=k}^{\infty} \beta_j} \right\} + \frac{|\beta_1 y_1^*|_A}{\sum_{j=2}^{\infty} \beta_j} \right) \\
&< \left(\sum_{j=n+1}^{\infty} \beta_j \right) \sum_{k=1}^{n} \left(\frac{\beta_k \alpha_k (1+\epsilon_k)}{\sum_{j=k+1}^{\infty} \beta_j \sum_{j=k}^{\infty} \beta_j} \right),
\end{aligned}$$

with the appropriate intermediate inequalities omitted if n is small. This same inequality also holds if $n = 1$ (with all intermediate inequalities omitted, of course) because it just reduces to the inequality $\beta_1 |y_1^*|_A < \beta_1\alpha_1(1+\epsilon_1)$. Since $\alpha_k \le \alpha$ for each k,

$$\begin{aligned}\left|\sum_{j=1}^{n}\beta_j y_j^*\right|_A &< \alpha\left(\sum_{j=n+1}^{\infty}\beta_j\right)\sum_{k=1}^{n}\left(\frac{\beta_k(1+\epsilon_k)}{\sum_{j=k+1}^{\infty}\beta_j\sum_{j=k}^{\infty}\beta_j}\right)\\ &< \alpha\left(\sum_{j=n+1}^{\infty}\beta_j\right)\left(\sum_{k=1}^{n}\left\{\frac{1}{\sum_{j=k+1}^{\infty}\beta_j}-\frac{1}{\sum_{j=k}^{\infty}\beta_j}\right\}+(1-\theta)\right)\\ &\doteq \alpha\left(1-\theta\sum_{j=n+1}^{\infty}\beta_j\right),\end{aligned}$$

which is the inequality needed in (c). ■

1.13.11 Theorem. (R. C. James, 1957 [111], 1972 [117]). *Let X be a separable Banach space. Then the following are equivalent.*

(a) *The space X is not reflexive.*

(b) *If $0 < \theta < 1$, then there is a sequence (x_n^*) in B_{X^*} such that $\lim_n x_n^* x = 0$ for each x in X and $d\big(0, \mathrm{co}(\{x_n^* : n \in \mathbb{N}\})\big) \ge \theta$.*

(c) *If $0 < \theta < 1$ and (β_n) is a sequence of positive numbers with sum 1, then there is an α such that $\theta \le \alpha \le 1$ and a sequence (y_n^*) in B_{X^*} such that*

   (1) *$\lim_n y_n^* x = 0$ for each x in X;*

   (2) *$\|\sum_{j=1}^{\infty}\beta_j y_j^*\| = \alpha$; and*

   (3) *$\|\sum_{j=1}^{n}\beta_j y_j^*\| < \alpha(1-\theta\sum_{j=n+1}^{\infty}\beta_j)$ for each positive integer n.*

(d) *There is a z^* in X^* that is not a norm-attaining functional.*

PROOF. To see that (a) ⇒ (b), suppose that X is not reflexive and that $0 < \theta < 1$. Let Q be the natural map from X into X^{**} and let x^{**} be an element of X^{**} such that

$$\theta < d\big(x^{**}, Q(X)\big) = \|x^{**} + Q(X)\| \le \|x^{**}\| \le 1.$$

Let $\{x_n : n \in \mathbb{N}\}$ be a countable dense subset of X. The immediate goal is to construct a sequence (x_n^*) in X^* such that, for each n,

(i) $\|x_n^*\| \le 1$;

(ii) $x^{**}x_n^* = \theta$; and

(iii) $x_n^* x_j = 0$ whenever $j \le n$.

It will then be shown that (x_n^*) satisfies the conclusion of (b).

Let $M = \theta/d(x^{**}, Q(X))$. Notice that $0 < M < 1$. For the remainder of this paragraph, let n be a fixed positive integer. If $\alpha_1, \dots, \alpha_n$ are scalars, then

$$M\left\|x^{**} + \sum_{j=1}^{n} \alpha_j Q x_j\right\| \ge M \cdot d(x^{**}, Q(X)) = \theta.$$

Let $c_j = 0$ when $j = 1, \dots, n$ and let $c_{n+1} = \theta$. Then for each linear combination $\alpha_1 Q x_1 + \cdots + \alpha_n Q x_n + \alpha_{n+1} x^{**}$ of $Qx_1, \dots, Qx_n, x^{**}$,

$$\left|\sum_{j=1}^{n+1} \alpha_j c_j\right| = |\alpha_{n+1}|\theta \le M\left\|\sum_{j=1}^{n} \alpha_j Q x_j + \alpha_{n+1} x^{**}\right\|.$$

By Helly's theorem, there is for each positive ϵ a y_ϵ^* in X^* such that

(iv) $\|y_\epsilon^*\| \le M + \epsilon$;

(v) $(Qx_j)(y_\epsilon^*) = c_j = 0$ when $j = 1, \dots, n$; and

(vi) $x^{**} y_\epsilon^* = c_{n+1} = \theta$.

Letting $x_n^* = y_\epsilon^*$ for a suitably small ϵ yields an x_n^* satisfying (i), (ii), and (iii).

If $x^* \in \operatorname{co}(\{x_n^* : n \in \mathbb{N}\})$, then the validity of (ii) for each n implies that $\|x^*\| \ge |x^{**} x^*| = \theta$, and so

$$d(0, \operatorname{co}(\{x_n^* : n \in \mathbb{N}\})) \ge \theta.$$

Suppose that $x_0 \in X$ and $k \in \mathbb{N}$. If $n \in \mathbb{N}$, then

$$|x_n^* x_0| \le |x_n^* x_k| + |x_n^*(x_0 - x_k)| \le |x_n^* x_k| + \|x_0 - x_k\|.$$

Since $\lim_n x_n^* x_k = 0$ and $\{x_j : j \in \mathbb{N}\}$ is dense in X, it follows that $\lim_n x_n^* x_0 = 0$. This finishes the proof that (a) $\Rightarrow$ (b).

Suppose now that (b) holds and that θ is a real number and (β_n) a sequence of positive real numbers such that $0 < \theta < 1$ and $\sum_n \beta_n = 1$. Let (x_n^*) be a sequence as in the conclusion of (b) and let (y_n^*) be the sequence in B_{X^*} and α the scalar such that $\theta \le \alpha \le 1$ guaranteed when Lemma 1.13.10 is applied with A equal to B_X. Then (y_n^*) and α do all that is required of them in (c). In particular, if $x \in X$ then the facts that $y_n^* \in \operatorname{co}(\{x_j^* : j \ge n\})$ for each n and $\lim_n x_n^* x = 0$ together imply that $\lim_n y_n^* x = 0$. Therefore (b) $\Rightarrow$ (c).

Suppose that (c) holds. Let θ be any scalar and (β_n) any sequence of positive scalars such that $0 < \theta < 1$ and $\sum_n \beta_n = 1$. Let α and (y_n^*) be as in (c) and let $z^* = \sum_{j=1}^{\infty} \beta_j y_j^*$. Then $\|z^*\| = \alpha$. It will be shown that z^* is not norm-attaining. Let x be an element of B_X and let n be a positive integer such that $|y_j^* x| < \alpha\theta$ whenever $j > n$. Then

$$\begin{aligned}
|z^* x| &= \left|\sum_{j=1}^{\infty} \beta_j y_j^* x\right| \\
&\le \left|\sum_{j=1}^{n} \beta_j y_j^* x\right| + \sum_{j=n+1}^{\infty} \beta_j |y_j^* x| \\
&< \left\|\sum_{j=1}^{n} \beta_j y_j^*\right\| + \alpha\theta \sum_{j=n+1}^{\infty} \beta_j \\
&< \alpha\left(1 - \theta \sum_{j=n+1}^{\infty} \beta_j\right) + \alpha\theta \sum_{j=n+1}^{\infty} \beta_j \\
&= \alpha \\
&= \|z^*\|,
\end{aligned}$$

so z^* does not attain its norm at x. This proves that (c) $\Rightarrow$ (d).

Finally, Proposition 1.11.11 assures that each bounded linear functional on a reflexive normed space is norm-attaining, so (d) $\Rightarrow$ (a). ■

Thus, a separable Banach space X is reflexive if and only if each x^* in X^* is a norm-attaining functional. To get the same result for arbitrary Banach spaces, some temporary notation is needed.

Let X be a *real* normed space and let (x_n^*) be a bounded sequence in X^*. Let

$$L(x_n^*) = \{\, x^* : x^* \in X^*,\ x^* x \le \limsup_n x_n^* x \text{ whenever } x \in X \,\}$$

and

$$V(x_n^*) = \{\, (y_n^*) : (y_n^*) \text{ is a sequence in } X^*, \text{ each } y_n^* \in \operatorname{co}(\{x_n^*, x_{n+1}^*, \dots\}) \,\}.$$

This notation is needed only for the next few results and does not apply outside this section, with the only exception being that the notation for $L(x_n^*)$ will be temporarily reinstated in Section 2.9. Notice that $(x_n^*) \in V(x_n^*)$, that each member of $V(x_n^*)$ is a sequence in the closed ball centered at 0 of radius $\sup\{\, \|x_n^*\| : n \in \mathbb{N} \,\}$, and that $V(y_n^*) \subseteq V(x_n^*)$ and $L(y_n^*) \subseteq L(x_n^*)$ whenever $(y_n^*) \in V(x_n^*)$. Notice also that if $x^* \in L(x_n^*)$, then $\|x^*\| \le \sup\{\, \|x_n^*\| : n \in \mathbb{N} \,\}$ since $|x^* x| \le \sup\{\, \|x_n^*\| : n \in \mathbb{N} \,\}\|x\|$ for each x in X, and $\liminf_n x_n^* x \le x^* x$ whenever $x \in X$ since $\liminf_n x_n^* x = -\limsup_n x_n^*(-x)$.

1.13.12 Lemma. *Let X be a real normed space and let (x_n^*) be a bounded sequence in X^*. Then $L(x_n^*)$ is nonempty.*

PROOF. Let $p(x) = \limsup_n x_n^* x$ for each x in X. Then p is a sublinear functional on X. Let y^* be the zero functional on the subspace $\{0\}$ of X.

Then $y^*(0) = p(0)$, so the vector space version of the Hahn-Banach extension theorem implies that y^* has a linear extension x^* to X such that $x^*x \le p(x)$ whenever $x \in X$. Since $|x^*x| \le \sup\{\,\|x_n^*\| : n \in \mathbb{N}\,\}\|x\|$ whenever $x \in X$, the linear functional x^* is bounded, and so $x^* \in L(x_n^*)$. ■

The following technical lemma is the heart of the proof of James's theorem for arbitrary Banach spaces. Notice its similarity to Lemma 1.13.10. As with Lemma 1.13.10, it is proved in a bit more generality than is needed in this section, but the extra generality will be required in Section 2.9.

1.13.13 Lemma. (R. C. James, 1972 [117]). *Let A be a nonempty balanced subset of the closed unit ball of a real normed space X. Suppose that (β_n) is a sequence of positive numbers with sum 1, that $0 < \theta < 1$, and that (x_n^*) is a sequence in B_{X^*} such that $\sup\{\,|(x^* - w^*)(x)| : x \in A\,\} \ge \theta$ whenever $x^* \in \operatorname{co}(\{\,x_n^* : n \in \mathbb{N}\,\})$ and $w^* \in L(x_n^*)$. Then there is an α such that $\theta \le \alpha \le 2$ and a sequence (y_n^*) in B_{X^*} such that whenever $w^* \in L(y_n^*)$,*

(a) $\sup\{\,|\sum_{j=1}^\infty \beta_j(y_j^* - w^*)(x)| : x \in A\,\} = \alpha$; *and*

(b) $\sup\{\,|\sum_{j=1}^n \beta_j(y_j^* - w^*)(x)| : x \in A\,\} < \alpha(1 - \theta\sum_{j=n+1}^\infty \beta_j)$ *for each positive integer n.*

PROOF. This proof consists primarily of the verification of a collection of eight claims, the first six of which appear in an induction argument. Lemma 1.13.12 and the remarks immediately preceding it are used extensively along the way.

As in the proof of Lemma 1.13.10, let $|x^*|_A = \sup\{\,|x^*x| : x \in A\,\}$ whenever $x^* \in X^*$. Then $|\cdot|_A$ is a continuous seminorm on X, and $|x^*|_A \le \|x^*\|$ for each x^* in X^*. Let (ϵ_n) be a sequence of positive reals converging to 0 such that

$$\sum_{k=1}^\infty \frac{\beta_k \epsilon_k}{\sum_{j=k+1}^\infty \beta_j \sum_{j=k}^\infty \beta_j} < 1 - \theta.$$

The first order of business is to use induction to obtain a sequence (α_j) of scalars and sequences $(y_j^*); ({}^0x_j^*), ({}^1x_j^*), ({}^2x_j^*), \dots ; ({}^1z_j^*), ({}^2z_j^*), ({}^3z_j^*), \dots$ in X^* such that $({}^0x_j^*)$ lies in B_{X^*} and, for each positive integer n,

(1) y_n^* and the sequences $({}^nz_j^*)$ and $({}^nx_j^*)$ lie in B_{X^*};

(2) $({}^nz_j^*) \in V({}^{n-1}x_j^*)$;

(3) $({}^nx_j^*)$ is a subsequence of $({}^nz_j^*)$;

(4) $y_n^* \in \operatorname{co}(\{{}^{n-1}x_n^*, {}^{n-1}x_{n+1}^*, {}^{n-1}x_{n+2}^*, \dots\})$;

(5) $\theta \le \alpha_n \le 2$; and

(6) $\alpha_{n-1} \le \alpha_n$ if $n \ge 2$.

To start the induction, let $({}^0x_j^*) = (x_j^*)$. Now suppose that $m \in \mathbb{N}$ and, if $m \ge 2$, that α_n, y_n^*, $({}^nz_j^*)$, and $({}^nx_j^*)$ have been chosen to satisfy (1) through (6) when $n = 1, \ldots, m-1$. In what follows, a sum from 1 to some upper limit N is to be considered to be 0 if $N \le 0$.

Claim 1: If $(v_j^*) \in V({}^{m-1}x_j^*)$, then (v_j^*) lies in B_{X^*}. To prove this, it is enough to show that $({}^{m-1}x_j^*)$ lies in B_{X^*}, which follows immediately from the definition of $({}^0x_j^*)$ if $m = 1$ or from the induction hypothesis if $m \ge 2$.

Claim 2: If $y^* \in \operatorname{co}(\{{}^{m-1}x_m^*, {}^{m-1}x_{m+1}^*, {}^{m-1}x_{m+2}^*, \ldots\})$ and $(v_j^*) \in V({}^{m-1}x_j^*)$, then the formula

$$S_m(y^*, (v_j^*)) = \left\{ \left| \sum_{j=1}^{m-1} \beta_j y_j^* + \left(\sum_{j=m}^{\infty} \beta_j \right) y^* - w^* \right|_A : w^* \in L(v_j^*) \right\}$$

defines a nonempty subset of $[0,2]$, so the following formula defines a number in $[0,2]$:

$$\alpha_m = \inf\{ \sup S_m(y^*, (v_j^*)) : \\ y^* \in \operatorname{co}(\{{}^{m-1}x_m^*, {}^{m-1}x_{m+1}^*, \ldots\}), (v_j^*) \in V({}^{m-1}x_j^*) \}.$$

To see this, fix a y^* in $\operatorname{co}(\{{}^{m-1}x_m^*, {}^{m-1}x_{m+1}^*, \ldots\})$ and a (v_j^*) in $V({}^{m-1}x_j^*)$. Then $y^* \in B_{X^*}$ since $({}^{m-1}x_j^*)$ lies in B_{X^*}, and $L(v_j^*) \subseteq B_{X^*}$ by Claim 1. Furthermore, if $m \ge 2$ then $y_1^*, \ldots, y_{m-1}^* \in B_{X^*}$ by the induction hypothesis. It is an easy consequence of all this that if $w^* \in L(v_j^*)$, then

$$0 \le \left| \sum_{j=1}^{m-1} \beta_j y_j^* + \left(\sum_{j=m}^{\infty} \beta_j \right) y^* - w^* \right|_A \le 2.$$

Since $L(v_j^*) \ne \varnothing$, Claim 2 follows.

Claim 3: If $m \ge 2$, then $V({}^{m-2}x_j^*) \supseteq V({}^{m-1}x_j^*)$. For this, just notice that if $m \ge 2$, then $({}^{m-1}x_j^*)$ is a subsequence of the element $({}^{m-1}z_j^*)$ of $V({}^{m-2}x_j^*)$, which implies that ${}^{m-1}x_j^* \in \operatorname{co}(\{{}^{m-2}x_j^*, {}^{m-2}x_{j+1}^*, \ldots\})$ for each j, that is, that $({}^{m-1}x_j^*) \in V({}^{m-2}x_j^*)$.

Claim 4: If $m \ge 2$ and $x^* \in \operatorname{co}(\{{}^{m-1}x_m^*, {}^{m-1}x_{m+1}^*, {}^{m-1}x_{m+2}^*, \ldots\})$, then

$$\frac{\beta_{m-1}}{\sum_{j=m-1}^{\infty} \beta_j} y_{m-1}^* + \frac{\sum_{j=m}^{\infty} \beta_j}{\sum_{j=m-1}^{\infty} \beta_j} x^* \\ \in \operatorname{co}(\{{}^{m-2}x_{m-1}^*, {}^{m-2}x_m^*, {}^{m-2}x_{m+1}^*, \ldots\}).$$

To see this, first notice that $({}^{m-1}x_j^*) \in V({}^{m-2}x_j^*)$ by Claim 3, and so ${}^{m-1}x_j^* \in \operatorname{co}(\{\, {}^{m-2}x_k^* : k \ge m-1 \,\})$ when $j \ge m$. It follows that $x^* \in \operatorname{co}(\{\, {}^{m-2}x_k^* : k \ge m-1 \,\})$. Since $y_{m-1}^* \in \operatorname{co}(\{\, {}^{m-2}x_k^* : k \ge m-1 \,\})$ by the induction hypothesis, every convex combination of y_{m-1}^* and x^* is in $\operatorname{co}(\{\, {}^{m-2}x_k^* : k \ge m-1 \,\})$, which proves Claim 4.

Claim 5: If $m \ge 2$, then $\alpha_{m-1} \le \alpha_m$. To see this, notice that because of Claims 3 and 4 and the fact that the infimum of a subset of a set is at least as large as the infimum of the entire set,

$$\begin{aligned}
\alpha_{m-1} &\le \inf\Big\{ \sup S_{m-1}(y^*, (v_j^*)) : y^* = \frac{\beta_{m-1}}{\sum_{j=m-1}^{\infty}\beta_j} y_{m-1}^* + \frac{\sum_{j=m}^{\infty}\beta_j}{\sum_{j=m-1}^{\infty}\beta_j} x^*, \\
&\qquad x^* \in \operatorname{co}(\{{}^{m-1}x_m^*, {}^{m-1}x_{m+1}^*, \dots\}), (v_j^*) \in V({}^{m-1}x_j^*) \Big\} \\
&= \inf\Big\{ \sup\Big\{ \Big| \sum_{j=1}^{m-2} \beta_j y_j^* + \beta_{m-1} y_{m-1}^* \\
&\qquad + \Big(\sum_{j=m}^{\infty}\beta_j\Big) x^* - w^* \Big|_A : w^* \in L(v_j^*) \Big\} : \\
&\qquad x^* \in \operatorname{co}(\{{}^{m-1}x_m^*, {}^{m-1}x_{m+1}^*, \dots\}), (v_j^*) \in V({}^{m-1}x_j^*) \Big\} \\
&= \inf\Big\{ \sup S_m(x^*, (v_j^*)) : \\
&\qquad x^* \in \operatorname{co}(\{{}^{m-1}x_m^*, {}^{m-1}x_{m+1}^*, \dots\}), (v_j^*) \in V({}^{m-1}x_j^*) \Big\} \\
&= \alpha_m,
\end{aligned}$$

as claimed.

Claim 6: $\theta \le \alpha_m \le 2$. To prove this, notice that because of Claims 2 and 5 it is enough to show that $\alpha_1 \ge \theta$. Let y^* be an element of $\operatorname{co}(\{x_j^* : j \in \mathbb{N}\})$ and (v_j^*) an element of $V(x_j^*)$. It is enough to show that $\sup S_1(y^*, (v_j^*)) \ge \theta$, that is, that $\sup\{ |y^* - w^*|_A : w^* \in L(v_j^*) \} \ge \theta$. Let w^* be an element of $L(v_j^*)$. It is enough to show that $|y^* - w^*|_A \ge \theta$. But this is true by the hypotheses of this lemma, since $L(v_j^*) \subseteq L(x_j^*)$, which proves Claim 6.

Using the definition of α_m, choose y_m^* from $\operatorname{co}(\{{}^{m-1}x_m^*, {}^{m-1}x_{m+1}^*, \dots\})$ and $({}^m z_j^*)$ from $V({}^{m-1}x_j^*)$ so that

$$\begin{aligned}
\alpha_m &\le \sup\Big\{ \Big| \sum_{j=1}^{m-1} \beta_j y_j^* + \Big(\sum_{j=m}^{\infty}\beta_j\Big) y_m^* - w^* \Big|_A : w^* \in L({}^m z_j^*) \Big\} \qquad (1.6) \\
&< \alpha_m(1+\epsilon_m),
\end{aligned}$$

then choose w_m^* from $L({}^m z_j^*)$ so that

$$\alpha_m(1-\epsilon_m) < \Big| \sum_{j=1}^{m-1} \beta_j y_j^* + \Big(\sum_{j=m}^{\infty}\beta_j\Big) y_m^* - w_m^* \Big|_A .$$

The fact that A is balanced assures that there is an x_m in A such that

$$\alpha_m(1-\epsilon_m) < \sum_{j=1}^{m-1} \beta_j y_j^* x_m + \Big(\sum_{j=m}^{\infty}\beta_j\Big) y_m^* x_m - w_m^* x_m. \qquad (1.7)$$

It follows from Claim 1 that $|\liminf_j({}^m z_j^* x_m)| \le 1$. Let $({}^m x_j^*)$ be a subsequence of $({}^m z_j^*)$ such that $\lim_j({}^m x_j^* x_m) = \liminf_j({}^m z_j^* x_m)$. It then follows from the fact that $({}^{m-1}x_j^*)$ lies in B_{X^*}, along with Claims 1, 5, and 6, that α_m, y_m^*, $({}^m z_j^*)$, and $({}^m x_j^*)$ do what is required of them in (1) through (6) when $n = m$. This completes the induction.

Claim 7: $L(y_j^*) \subseteq \bigcap_{n=0}^\infty L({}^n x_j^*) \subseteq \bigcap_{n=1}^\infty L({}^n z_j^*)$. To see that the second inclusion holds, suppose that $n \in \mathbb{N}$. It is enough to show that $L({}^n x_j^*) \subseteq L({}^n z_j^*)$, for which it is enough to show that $({}^n x_j^*) \in V({}^n z_j^*)$, and this follows from the fact that $({}^n x_j^*)$ is a subsequence of $({}^n z_j^*)$. For the proof that $L(y_j^*) \subseteq \bigcap_{n=0}^\infty L({}^n x_j^*)$, suppose that n is a nonnegative integer. Notice that if $j \in \mathbb{N}$, then (2), (3), and (4) together imply that

$$\begin{aligned} y_j^* &\in \operatorname{co}(\{{}^{j-1}x_j^*, {}^{j-1}x_{j+1}^*, \dots\}) \subseteq \operatorname{co}(\{{}^{j-1}z_j^*, {}^{j-1}z_{j+1}^*, \dots\}) \\ &\subseteq \operatorname{co}(\{{}^{j-2}x_j^*, {}^{j-2}x_{j+1}^*, \dots\}) \subseteq \operatorname{co}(\{{}^{j-2}z_j^*, {}^{j-2}z_{j+1}^*, \dots\}) \\ &\subseteq \quad \cdots \\ &\subseteq \operatorname{co}(\{{}^0x_j^*, {}^0x_{j+1}^*, \dots\}). \end{aligned}$$

Therefore $y_j^* \in \operatorname{co}(\{{}^n x_j^*, {}^n x_{j+1}^*, \dots\})$ when $j > n$, from which it follows that $\limsup_j(y_j^* x) \le \limsup_j({}^n x_j^* x)$ for each x in X and thus that $L(y_j^*) \subseteq L({}^n x_j^*)$. This finishes the proof of Claim 7.

Claim 8: If $w^* \in L(y_j^*)$ and $m \in \mathbb{N}$, then (1.7) holds for the same element x_m of A when w_m^* is replaced by w^*. To prove this, notice from Claim 7 that $w^* \in L({}^m x_j^*)$, and so the way that $({}^m x_j^*)$ was obtained from $({}^m z_j^*)$ assures that

$$w^* x_m = \lim_j({}^m x_j^* x_m) = \liminf_j({}^m z_j^* x_m) \le w_m^* x_m.$$

Claim 8 follows from this.

Fix a w^* in $L(y_j^*)$. It follows from (1.6) and Claims 7 and 8 that if $n \in \mathbb{N}$, then

$$\alpha_n(1-\epsilon_n) < \left|\sum_{j=1}^{n-1} \beta_j y_j^* + \left(\sum_{j=n}^\infty \beta_j\right) y_n^* - w^*\right|_A < \alpha_n(1+\epsilon_n). \qquad (1.8)$$

Claims 5 and 6 together assure that $\lim_n \alpha_n$ exists and lies in $[\theta, 2]$. Call this limit α. Taking limits as n tends to infinity in (1.8) shows that

$$\alpha = \left|\sum_{j=1}^\infty \beta_j(y_j^* - w^*)\right|_A.$$

All that remains to be proved is that

$$\left|\sum_{j=1}^n \beta_j(y_j^* - w^*)\right|_A < \alpha\left(1 - \theta\sum_{j=n+1}^\infty \beta_j\right)$$

when $n \in \mathbb{N}$. The inequalities needed to do this are exactly those that appear in step 6 of the proof of Lemma 1.13.10, except that y_k^* must be replaced by $(y_k^* - w^*)$ for each k. ■

1.13.14 Theorem. (R. C. James, 1964 [112], 1972 [117]). *Let X be a real Banach space. Then the following are equivalent.*

(a) *The space X is not reflexive.*

(b) *If $0 < \theta < 1$, then there is a closed subspace M of X and a sequence (x_n^*) in B_{X^*} such that $d(M^\perp, \operatorname{co}(\{x_n^* : n \in \mathbb{N}\})) \geq \theta$ and $\lim_n x_n^* x = 0$ for each x in M.*

(c) *If $0 < \theta < 1$ and (β_n) is a sequence of positive numbers with sum 1, then there is an α such that $\theta \leq \alpha \leq 2$ and a sequence (y_n^*) in B_{X^*} such that whenever $w^* \in L(y_n^*)$,*

  (1) $\|\sum_{j=1}^{\infty} \beta_j (y_j^* - w^*)\| = \alpha$; *and*

  (2) $\|\sum_{j=1}^{n} \beta_j (y_j^* - w^*)\| < \alpha(1 - \theta \sum_{j=n+1}^{\infty} \beta_j)$ *for each positive integer n.*

(d) *There is a z^* in X^* that is not a norm-attaining functional.*

PROOF. For the proof that (a) ⇒ (b), suppose that X is not reflexive and that $0 < \theta < 1$. It follows from Theorem 1.13.8 that some separable closed subspace M of X is not reflexive. By Theorem 1.13.11, there is a sequence (m_n^*) in B_{M^*} such that $d(0, \operatorname{co}(\{m_n^* : n \in \mathbb{N}\})) \geq \theta$ and $\lim_n m_n^* x = 0$ for each x in M. For each positive integer n, let x_n^* be a Hahn-Banach extension of m_n^* to X. If $x^* \in \operatorname{co}(\{x_n^* : n \in \mathbb{N}\})$ and $y^* \in M^\perp$, then the restriction of $x^* - y^*$ to M is a member m^* of $\operatorname{co}(\{m_n^* : n \in \mathbb{N}\})$, and so

$$\begin{aligned}\|x^* - y^*\| &\geq \sup\{\, |(x^* - y^*)(m)| : m \in B_M \,\}\\ &= \|m^*\|\\ &\geq d(0, \operatorname{co}(\{m_n^* : n \in \mathbb{N}\}))\\ &\geq \theta.\end{aligned}$$

It follows that

$$d(M^\perp, \operatorname{co}(\{x_n^* : n \in \mathbb{N}\})) \geq \theta,$$

so (x_n^*) is a sequence as in the conclusion of (b). Therefore (a) ⇒ (b).

Suppose that (b) holds. Let θ be a scalar and (β_n) a sequence of positive scalars such that $0 < \theta < 1$ and $\sum_n \beta_n = 1$. For this θ, let M be a closed subspace of X and (x_n^*) a sequence in B_{X^*} with the properties guaranteed by (b). It is easy to check that $L(x_n^*) \subseteq M^\perp$, so

$$d(L(x_n^*), \operatorname{co}(\{x_n^* : n \in \mathbb{N}\})) \geq \theta.$$

Letting A equal B_X in Lemma 1.13.13 yields an α in $[\theta, 2]$ and a sequence (y_n^*) in B_{X^*} satisfying (1) and (2) in (c), which proves that (b) $\Rightarrow$ (c).

Suppose that (c) holds. Let θ and Δ be scalars such that $0 < \theta < 1$ and $0 < \Delta < \theta^2/2$. For each positive integer n, let

$$\beta_n = \frac{2-\Delta}{\Delta}\left(\frac{\Delta}{2}\right)^n.$$

Then (β_n) is a sequence of positive scalars that sums to 1. Let α and (y_n^*) be as in (c). Let w^* be any member of $L(y_n^*)$ and let $z^* = \sum_{j=1}^\infty \beta_j(y_j^* - w^*)$. Then $\|z^*\| = \alpha$. It will be shown that z^* is not norm-attaining. Suppose that $x \in B_X$. Since $\liminf_j y_j^* x \le w^* x$ and $\theta \le \alpha$, there is a positive integer n such that

$$(y_{n+1}^* - w^*)(x) < \theta^2 - 2\Delta \le \alpha\theta - 2\Delta.$$

Since $w^* y \le \limsup_j y_j^* y \le 1$ whenever $y \in B_X$, it follows that $\|w^*\| \le 1$. Therefore

$$\begin{aligned}
z^*x &= \sum_{j=1}^\infty \beta_j(y_j^* - w^*)(x)\\
&< \sum_{j=1}^n \beta_j(y_j^* - w^*)(x) + (\alpha\theta - 2\Delta)\beta_{n+1} + \sum_{j=n+2}^\infty \beta_j(y_j^* - w^*)(x)\\
&\le \left\|\sum_{j=1}^n \beta_j(y_j^* - w^*)\right\| + (\alpha\theta - 2\Delta)\beta_{n+1} + 2\sum_{j=n+2}^\infty \beta_j\\
&< \alpha\left(1 - \theta\sum_{j=n+1}^\infty \beta_j\right) + (\alpha\theta - 2\Delta)\beta_{n+1} + 2\sum_{j=n+2}^\infty \beta_j.
\end{aligned}$$

Since $\sum_{j=n+2}^\infty \beta_j = \frac{1}{2}\Delta\sum_{j=n+1}^\infty \beta_j < \Delta\sum_{j=n+1}^\infty \beta_j$,

$$\begin{aligned}
z^*x &< \alpha - (\alpha\theta - 2\Delta)\sum_{j=n+1}^\infty \beta_j + (\alpha\theta - 2\Delta)\beta_{n+1}\\
&= \alpha - (\alpha\theta - 2\Delta)\sum_{j=n+2}^\infty \beta_j\\
&< \alpha\\
&= \|z^*\|.
\end{aligned}$$

Since $-x \in B_X$, it also follows that $-z^*x = z^*(-x) < \|z^*\|$, so $|z^*x| < \|z^*\|$. Therefore z^* is not norm-attaining, which proves that (c) $\Rightarrow$ (d).

Finally, Proposition 1.11.11 assures that each bounded linear functional on a reflexive normed space is norm-attaining, so (d) $\Rightarrow$ (a). ■

The big theorem of this section is now an easy corollary of the result just proved.

1.13.15 James's Theorem. (R. C. James, 1964 [112]). *If every bounded linear functional on a Banach space is norm-attaining, then the space is reflexive.*

PROOF. Let X be a Banach space such that every bounded linear functional on X is norm-attaining. If X is a real Banach space, then the preceding theorem yields the desired result, so it may be assumed that X is a complex Banach space. Let X_r be the real Banach space obtained from X by using only real scalars. Let u^* be a bounded linear functional on X_r and let x^* be the member of X^* such that $\operatorname{Re} x^* = u^*$. Then there is an x in B_X and a scalar α with modulus 1 such that

$$\|u^*\| = \|x^*\| = |x^*x| = x^*(\alpha x) = u^*(\alpha x).$$

Since each member of $(X_r)^*$ is norm-attaining, the space X_r is reflexive, and so X is reflexive. ∎

In 1971, James [116] gave an example to show that an incomplete normed space can have the property that all of its bounded linear functionals are norm-attaining, so the completeness hypothesis in the preceding theorem cannot in general be omitted. However, see Exercise 1.150.

Here is a summary of the most important results of this section along with a few gleaned from Section 1.11. Though this summary is stated only for Banach spaces, conditions (a), (c), (d), (e), and (f) are equivalent without the completeness hypothesis, since each implies completeness; see the results from which the equivalences were taken. Notice that the fact that (a) implies (f) and (g) follows trivially by considering the identity operator on the space.

1.13.16 Summary. *Suppose that X is a Banach space. Then the following are equivalent.*

(a) *The space X is reflexive.*
(b) *The dual space of X is reflexive.*
(c) *Each bounded sequence in X has a weakly convergent subsequence.*
(d) *Whenever (C_n) is a sequence of nonempty closed bounded convex sets in X such that $C_n \supseteq C_{n+1}$ for each n, it follows that $\bigcap_n C_n \neq \varnothing$.*
(e) *Each separable closed subspace of X is reflexive.*
(f) *The space X is isomorphic to a reflexive space.*
(g) *There is a bounded linear operator from some reflexive space onto X.*
(h) *Each bounded linear functional on X is norm-attaining.*

(i) *The following does* not *hold: For each θ such that $0 < \theta < 1$ there is a sequence (x_n^*) in S_{X^*} and a sequence (x_n) in S_X such that* $\operatorname{Re} x_n^* x_j \geq \theta$ *if $n \leq j$ and* $\operatorname{Re} x_n^* x_j = 0$ *if $n > j$.*

(j) *The following does* not *hold: For some θ such that $0 < \theta < 1$ there is a sequence (x_n^*) in S_{X^*} and a sequence (x_n) in S_X as in* (i).

(k) *The following does* not *hold: For each θ such that $0 < \theta < 1$ there is a sequence (x_n) in S_X such that*

$$d(\operatorname{co}(\{x_1, \ldots, x_n\}), \operatorname{co}(\{x_{n+1}, x_{n+2}, \ldots\})) \geq \theta$$

for each n.

(l) *The following does* not *hold: For some θ such that $0 < \theta < 1$ there is a sequence (x_n) in S_X as in* (k).

Another proof of the equivalence of (a), (c), and (e) will be obtained in Section 2.8 from the Eberlein-Šmulian theorem.

Exercises

1.146 Use an argument based on Proposition 1.13.1 to show that the nonreflexivity of complex c_0 follows from that of real c_0. (This is a trick question. To avoid the trap, think carefully about what is obtained from complex c_0 by restricting multiplication of vectors by scalars to $\mathbb{R} \times c_0$.)

1.147 Find a bounded sequence in ℓ_1 with no weakly convergent subsequence.

1.148 Find a sequence (C_n) of nonempty closed convex subsets of S_{ℓ_∞} such that $C_n \supseteq C_{n+1}$ for each n and $\bigcap_n C_n = \emptyset$.

1.149 In theory, Theorem 1.13.6 gives a test for the reflexivity of a normed space requiring no knowledge of its dual. In practice, it often happens that information obtained from the dual space is used to show that Theorem 1.13.6 can be applied. This is true in the proof of the nonreflexivity of $A(\mathbb{D})$ given in Example 1.13.7. In that proof, what information comes from $A(\mathbb{D})^*$?

1.150 Suppose that X is an incomplete normed space such that every element of X^* is norm-attaining. Prove that the completion of X is reflexive.

1.151 Suppose that C is a closed convex subset of a normed space X. Show that C is "weakly sequentially closed"; that is, that whenever (x_n) is a sequence in C that converges weakly to some element x of X, it follows that $x \in C$.

1.152 Prove that a Banach space is reflexive if and only if it has this property: Whenever A and B are nonempty closed convex subsets of the space such that $d(A, B) = 0$ and at least one of the two sets is bounded, the two sets intersect. Exercise 1.151 might help.

1.153 Let X be a Banach space. Prove that the following are equivalent.

(a) The space X is reflexive.

(b) Whenever C is a nonempty closed convex subset of X, there is a point of C nearest the origin; that is, there is at least one x in C such that $\|x\| = d(0, C)$.

(c) Whenever Y is a closed separable subspace of X and $y^* \in Y^*$, there is a point of $\{ y : y \in Y, \, y^*y = \|y^*\| \}$ nearest the origin.

If James's theorem is necessary in your argument, use only the version for separable spaces. If it is not, please send a copy of your proof to the author of this book!

1.154 Show that James's theorem follows from the form of James's weak compactness theorem stated in the discussion preceding Lemma 1.13.10.

1.155 Give a single example that shows that the conclusions of Theorems 1.13.4 and 1.13.9 do not necessarily hold if the normed space X in the statements of those theorems is not required to be complete.

1.156 This exercise uses the result of Exercise 1.151. Theorem 1.13.8 might lead one to wonder if a normed space must be separable whenever each of its reflexive subspaces is separable. The purpose of this exercise is to disprove this conjecture by showing that each reflexive subspace of ℓ_∞ is separable. Let X be a reflexive subspace of ℓ_∞. Justify each of the following statements.

(a) Let $\|(x_j)\|_a = \sum_j 2^{-j}|x_j|$ whenever $(x_j) \in X$. Then $\|\cdot\|_a$ is a norm on X. Furthermore, if $(x_j) \in X$ and $(x_j^{(n)})_{n=1}^\infty$ is a sequence in X bounded under the ℓ_∞ norm, then $\lim_n \|(x_j^{(n)}) - (x_j)\|_a = 0$ if and only if $\lim_n x_j^{(n)} = x_j$ for each j.

(b) The set B_X (meaning $B_{(X, \|\cdot\|_\infty)}$, not $B_{(X, \|\cdot\|_a)}$) is a compact, hence separable, metric space under the metric induced by $\|\cdot\|_a$. Let D be a countable subset of B_X dense in B_X with respect to this metric.

For the rest of this exercise, the normed space language and symbols refer to the space $(X, \|\cdot\|_\infty)$, not $(X, \|\cdot\|_a)$.

(c) For each x in B_X there is a sequence from D converging weakly to x.

(d) It follows that $[D] = X$, so X is separable.

2

The Weak and Weak* Topologies

The topology induced by a norm on a vector space is a very strong topology in the sense that it has many open sets. This has some advantages, especially since a function whose domain is such a space finds it particularly easy to be continuous, but it also has its disadvantages. For example, an infinite-dimensional normed space always has so many open sets that its closed unit ball cannot be compact. Because of this, many familiar facts about finite-dimensional normed spaces that are based on the Heine-Borel property cannot be immediately generalized to the infinite-dimensional case.

The main purpose of this chapter is to study topologies for normed spaces that are in general weaker than the norm topology, in the sense that they have fewer open sets, but that are still strong enough to have useful properties. The most important example of such a topology is the *weak* topology of a normed space X, which is the weakest topology for X such that every member of X^* is still continuous. Another useful topology is the *weak** (pronounced "weak star") topology of X^*. If Q is the natural map from X into X^{**}, then the weak* topology of X^* is the weakest topology for X^* such that every member of $Q(X)$ remains continuous. The primary emphasis of this chapter is on these two topologies, though it will often be convenient to carry out this study in a somewhat more general setting.

The topologies that will be discussed are not always induced by metrics, so familiar metric space arguments based on the convergence of sequences cannot be used in their usual form. However, most of those arguments can be adapted to general topological spaces if sequences are replaced by more general objects called *nets,* whose behavior is much like that of sequences. This chapter begins with a discussion of them.

2.1 Topology and Nets

For reference, and to make sure that the author and the reader are speaking the same topological language, here is a collection of the basic definitions from topology that will be needed. Many of these terms have already been encountered in Chapter 1 in a metric space setting.

2.1.1 Definition. Let X be a set. A *topology* for X is a collection $\mathfrak{T}$ of subsets of X such that

(1) both X and the empty set belong to $\mathfrak{T}$;

(2) for every subcollection of $\mathfrak{T}$, the union of the elements of the subcollection also belongs to $\mathfrak{T}$;

(3) for every finite subcollection of $\mathfrak{T}$, the intersection of the elements of the subcollection also belongs to $\mathfrak{T}$.

The set X with the topology $\mathfrak{T}$ is called the *topological space* $(X, \mathfrak{T})$, or just the *topological space* X when no confusion can result. The elements of $\mathfrak{T}$ are called *open sets.*

In the preceding definition, property (1) is actually redundant, since (2) and (3) assure that the union and intersection of the empty subcollection of $\mathfrak{T}$ are both in $\mathfrak{T}$.

2.1.2 Definition. Let $(X, \mathfrak{T})$ be a topological space and let x be an element of X. A *(local) basis for* $\mathfrak{T}$ *at* x is a collection $\mathfrak{B}_x$ of open sets containing x such that every open set containing x includes a member of $\mathfrak{B}_x$.

2.1.3 Definition. Let X be a set and let $\mathfrak{B}$ be a collection of subsets of X such that

(1) $\bigcup\{B : B \in \mathfrak{B}\} = X$;

(2) if $B_1, B_2 \in \mathfrak{B}$ and $x \in B_1 \cap B_2$, then there is a B_3 in $\mathfrak{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

Let $\mathfrak{T}$ be the collection of all sets that are unions of subcollections of $\mathfrak{B}$. Then $\mathfrak{T}$ is the *topology generated by the basis* $\mathfrak{B}$.

2.1.4 Definition. Let X be a set and let $\mathfrak{S}$ (Fraktur S; notice its resemblance to σ) be a collection of subsets of X. Let $\mathfrak{B}_{\mathfrak{S}}$ be the collection of all sets that are intersections of finitely many members of $\mathfrak{S}$. Then the *topology generated by the subbasis* $\mathfrak{S}$ is the topology generated by the basis $\mathfrak{B}_{\mathfrak{S}}$.

It is easy to check that the collection $\mathfrak{T}$ in Definition 2.1.3 really is a topology and that the collection $\mathfrak{B}_{\mathfrak{S}}$ in Definition 2.1.4 is a basis for a topology; for this latter fact, notice that the intersection X of the empty subcollection of $\mathfrak{S}$ is in $\mathfrak{B}_{\mathfrak{S}}$.

The following properties of bases and subbases for topologies follow directly from the above definitions. Every basis for a topology is also a subbasis for that topology. Every member of a basis or subbasis for a topology belongs to that topology. If $\mathfrak{T}$ is a topology for a set X, then a collection $\mathfrak{B}$ of subsets of X is a basis for $\mathfrak{T}$ if and only if every member of $\mathfrak{B}$ is open and every open set is a union of members of $\mathfrak{B}$, which happens if and only if the following holds: For each x in X, the family $\{ B : B \in \mathfrak{B}, x \in B \}$ is a basis for $\mathfrak{T}$ at x. If $\mathfrak{S}$ is a subbasis for a topology $\mathfrak{T}_{\mathfrak{S}}$ for a set X and $\mathfrak{T}$ is another topology for X such that $\mathfrak{S} \subseteq \mathfrak{T}$, then $\mathfrak{T}_{\mathfrak{S}} \subseteq \mathfrak{T}$; that is, the topology $\mathfrak{T}_{\mathfrak{S}}$ is the smallest topology for X that includes $\mathfrak{S}$.

2.1.5 Definition. Let $\{ X_\alpha : \alpha \in I \}$ be a family of topological spaces. Let $\mathfrak{S}$ be the collection of all subsets of the Cartesian product $\prod_{\alpha \in I} X_\alpha$ of the form $\prod_{\alpha \in I} U_\alpha$, where each U_α is open and at most one U_α is not equal to the corresponding X_α. Then the *product topology* of $\prod_{\alpha \in I} X_\alpha$ is the topology generated by the subbasis $\mathfrak{S}$. The *topological product* of the family of topological spaces is the Cartesian product with the product topology.

Notice that the basis generated by the subbasis in the preceding definition consists of all sets of the form $\prod_{\alpha \in I} U_\alpha$, where each U_α is open and $\{ \alpha : \alpha \in I, U_\alpha \neq X_\alpha \}$ is finite. *Henceforth, when the Cartesian product of topological spaces is treated as a topological space without the topology being specified, the product topology is implied.*

It is worth noting what happens when the index set I in the preceding definition is empty. In that case, the Cartesian product $\prod_{\alpha \in I} X_\alpha$ has as its lone element the empty set, viewed as a function from I into $\bigcup_{\alpha \in I} X_\alpha$; see, for example, [65, p. 22]. If $\mathfrak{S}$ and $\mathfrak{B}$ are, respectively, the subbasis and basis for the product topology of $\prod_{\alpha \in I} X_\alpha$ discussed above, then it is easy to check that $\mathfrak{S} = \mathfrak{B} = \{\prod_{\alpha \in I} X_\alpha\} = \{\{\emptyset\}\}$.

2.1.6 Definitions. Let X be a topological space.

(a) A subset of X is *closed* if its complement is open.

(b) A *neighborhood* of a point x in X is an open set that contains x.

(c) The space X is a T_0 space if, for each pair of distinct points in X, at least one has a neighborhood not containing the other.

(d) The space X is a T_1 space if, for each pair of distinct points in X, each has a neighborhood not containing the other.

(e) The space X is a *Hausdorff* or *separated* or T_2 space if, for each pair of distinct points x and y in X, there are disjoint neighborhoods U_x and U_y of x and y respectively.

(f) The space X is a *regular* or T_3 space if it is a T_1 space, and for each x in X and each closed subset F of X not containing x there are disjoint open sets U and V such that U is a neighborhood of x and V includes F.

(g) The space X is a *completely regular* or *Tychonoff* or $T_{3\frac{1}{2}}$ space if it is a T_1 space, and for each x in X and each closed subset F of X not containing x there is a continuous[1] function $f\colon X \to [0,1]$ such that $f(x) = 0$ and $f(y) = 1$ for each y in F.

(h) The space X is a *normal* or T_4 space if it is a T_1 space, and for each pair of disjoint closed subsets F_1 and F_2 of X there are disjoint open sets U_1 and U_2 that include F_1 and F_2 respectively.

(i) The *relative* or *induced* or *inherited* topology of a subset S of X is the collection of all sets $S \cap U$ such that U is open in X.

(j) The *closure* of a subset S of X, denoted by $\overline{S}$, is the smallest closed set that includes S, that is, the intersection of all closed sets that include S.

(k) The *interior* of a subset S of X, denoted by S°, is the largest open subset of S, that is, the union of all open subsets of S.

(l) The *boundary* of a subset S of X, denoted by ∂S, is the set $\overline{S} \cap \overline{X \setminus S}$, that is, the set $\overline{S} \setminus S^\circ$.

(m) A subset D of X is *dense* in another subset S of X if $D \subseteq S \subseteq \overline{D}$.

(n) A *limit point* or *cluster point* or *accumulation point* of a subset S of X is a point x in X such that each neighborhood of x contains at least one point of S distinct from x, that is, such that $x \in \overline{S \setminus \{x\}}$.

(o) A subset S of X is *compact* if, for each collection $\mathfrak{O}$ of open sets whose union includes S, there is a finite subcollection of $\mathfrak{O}$ whose union includes S. That is, the set S is compact if each open covering of S can be thinned to a finite subcovering. The set S is *relatively compact* if its closure is compact.

(p) The space X is *locally compact* if, for each x in X, there is a compact subset K_x of X such that $x \in K_x^\circ$.

(q) A subset S of X is *countably compact* if each countable open covering of S can be thinned to a finite subcovering. The set S is *relatively countably compact* if its closure is countably compact.

[1]The sharp-eyed reader will notice that continuity is mentioned here even though its definition does not appear until later. It is assumed that the reader is already familiar with continuity in topological spaces, as well as with most of the other terms cataloged in the early part of this section for reference.

(r) A subset S of X is *limit point compact* or *Fréchet compact* or has the *Bolzano-Weierstrass property* if each infinite subset of S has a limit point in S. The set S is *relatively limit point compact* if it satisfies the same condition except that the limit point need not be in S.

(s) A sequence (x_n) in X *converges* to an element x of X, and x is called a *limit* of (x_n), if, for each neighborhood U of x, there is a positive integer n_U such that $x_n \in U$ whenever $n \geq n_U$. This is denoted by writing $x_n \to x$ or $\lim_n x_n = x$.[2]

(t) A subset S of X is *sequentially compact* if each sequence in S has a convergent subsequence with a limit in S. The set S is *relatively sequentially compact* if it satisfies the same condition except that the limit need not be in S.

(u) The *sequential closure* of a subset S of X is the collection of all elements of X that are limits of sequences whose terms come from S. The set S is *sequentially closed* if it equals its sequential closure.

(v) A subset D of X is *sequentially dense* in another subset S of X if $D \subseteq S \subseteq \overline{D}^s$, where $\overline{D}^s$ is the sequential closure of D.

The conditions defined in (c) through (h) are sometimes called *separation axioms*. It is not hard to show that a topological space is a T_1 space if and only if each of its one-element subsets is closed; see Exercise 2.2 (a). From that and an application of Urysohn's lemma, it follows easily that $T_4 \Rightarrow T_{3\frac{1}{2}} \Rightarrow T_3 \Rightarrow T_2 \Rightarrow T_1 \Rightarrow T_0$.

Here are some relationships between the various types of compactness mentioned above. Most general topology texts, such as [65] or [172], will have proofs of the less obvious ones. In a metric space, the properties of compactness, countable compactness, limit point compactness, and sequential compactness are equivalent, as are the corresponding relative properties. Compactness obviously implies countable compactness. Countable compactness implies limit point compactness; the properties are equivalent in Hausdorff spaces. Even in Hausdorff spaces it is not always true that a set is relatively sequentially compact exactly when its closure is sequentially compact, or that a set is relatively limit point compact exactly when its closure is limit point compact; see Exercise 2.15. However, the equivalences do hold in metric spaces; see Exercise 2.1.

---

[2] Since two quantities equal to the same quantity should be equal to each other, it is best to avoid the notation $\lim_n x_n = x$ when (x_n) might have more than one limit. A similar comment applies to the notation for net convergence that will be introduced in Definition 2.1.14. This problem does not arise in Hausdorff spaces; see Proposition 2.1.17.

2.1.7 Definition. Let X and Y be topological spaces and let f be a function from X into Y. Then f is

(a) *continuous at the point* x_0 *of* X if, for each neighborhood V of $f(x_0)$, there is a neighborhood U of x_0 such that $f(U) \subseteq V$;

(b) *continuous* if, for each open subset V of Y, the set $f^{-1}(V)$ is open;

(c) *sequentially continuous at the point* x_0 *of* X if, whenever a sequence (x_n) in X converges to x_0, the sequence $\big(f(x_n)\big)$ converges to $f(x_0)$;

(d) *sequentially continuous* if, whenever a sequence (x_n) in X converges to a point x, the sequence $\big(f(x_n)\big)$ converges to $f(x)$;

(e) *open* if, for each open subset U of X, the set $f(U)$ is open;

(f) a *homeomorphism* if it is a bijection of X onto Y such that both f and f^{-1} are continuous; that is, if f is one-to-one, onto Y, continuous, and open.

A function from one topological space into another is continuous on its domain if and only if it is continuous at each point of its domain, and is sequentially continuous on its domain if and only if it is sequentially continuous at each point of its domain. Continuity at a point implies sequential continuity at that point, and the properties are equivalent when the domain is a metric space. Thus, global continuity implies global sequential continuity, with the properties equivalent when the domain is a metric space.

These continuity results for metric spaces can be generalized a bit. Recall that a topological space satisfies the *first countability axiom* if at each point of the space there is a countable basis for the topology at that point. Every metric space satisfies the first countability axiom; consider sequences of open balls with radii decreasing to 0. Now suppose that X is any topological space satisfying the first countability axiom. Then a function from X into a topological space is continuous at a point if and only if it is sequentially continuous at that point, so global continuity and global sequential continuity are also equivalent for such a function. Moreover, a subset of X is closed if and only if it is sequentially closed. The proofs are essentially the same as those for metric spaces; see, for example, [172, p. 190].

Not every topology permits such straightforward sequential testing for continuity and closure. For example, some topological spaces have subsets that are sequentially closed but not closed; a Hausdorff space with this property is constructed in Exercise 2.3. Though it might seem that sequential methods useful in metric spaces must be abandoned when working with topologies of this sort, many of those methods extend with very little modification to all topological spaces if sequences are replaced by *nets*.

2.1.8 Definition. A *directed set* is a nonempty set I with a relation $\preceq$ such that

(1) $\alpha \preceq \alpha$ whenever $\alpha \in I$;

(2) if $\alpha \preceq \beta$ and $\beta \preceq \gamma$, then $\alpha \preceq \gamma$;

(3) for each pair α, β of elements of I there is a $\gamma_{\alpha,\beta}$ in I such that $\alpha \preceq \gamma_{\alpha,\beta}$ and $\beta \preceq \gamma_{\alpha,\beta}$.

That is, a directed set is a nonempty preordered set that satisfies (3). A *net* or *Moore-Smith sequence* in a set X is a function from a directed set I into X. The set I is the *index set* for the net.

Some sources require that the preorder in the preceding definition be a partial order; that is, that the following additional requirement be included in the definition:

(4) $\alpha = \beta$ whenever $\alpha \preceq \beta$ and $\beta \preceq \alpha$.

The theory of nets can be developed either with or without this additional axiom. See [172, pp. 187–188] for a development that uses it.

The reason that the term *Moore-Smith sequence* is sometimes used for a net is that E. H. Moore and H. L. Smith [171] introduced nets in 1922 as the basis for a general theory of limits. Mauro Picone [190] devised the same theory independently in a book that appeared the next year. The term *net* was actually first used by J. L. Kelley [131] in a 1950 paper on topological convergence.[3] More on the early history of nets and the general theory of limits can be found in the survey articles by E. J. McShane [167] and R G. Bartle [18].

If $f\colon I \to X$ is a net, then for each α in I the α^{th} *term* $f(\alpha)$ of the net is often denoted by x_α, and the entire net is often denoted by $(x_\alpha)_{\alpha\in I}$ or just (x_α). By analogy with sequences, it is said that x_α *precedes* x_β in a net when $\alpha \preceq \beta$. In general, the familiar language of sequences is extended to nets whenever the meaning is clear.

2.1.9 Example. Every sequence is a net, with the directed set being $\mathbb{N}$ in its natural order.

2.1.10 Example. The set $\mathbb{R}$ with its natural order is a directed set, so this order makes every function with domain $\mathbb{R}$ into a net. Notice that a term in a net can be preceded by infinitely many others and that nets need not have first terms.

---

[3] The terminology was not Kelley's invention, though. Kelley had wanted to call such an object a *way*. However, nets have subnets, which Kelley would have dubbed *subways*. Norman Steenrod talked him out of it. After some prodding by Kelley, Steenrod suggested the term *net* as a substitute for *way*. See [204].

2.1.11 Example. The set $\mathbb{R}^2$ can be made into a directed set by declaring that $(\alpha_1, \beta_1) \preceq (\alpha_2, \beta_2)$ whenever $\alpha_1 \le \alpha_2$. If $x_{(\alpha,\beta)} = \alpha + \beta$ for each (α, β) in $\mathbb{R}^2$, then $(x_{(\alpha,\beta)})$ is a net in $\mathbb{R}$. Notice that $(1,2) \preceq (1,3)$ and $(1,3) \preceq (1,2)$ even though $(1,2) \neq (1,3)$.

2.1.12 Example. Let I be a three-element set $\{u, v, w\}$. Define $\preceq$ on I by letting these be all of the corresponding relations: $\alpha \preceq \alpha$ for each α in I; $u \preceq w$; and $v \preceq w$. This relation makes I into a directed set. Define a net (x_α) in $\mathbb{R}$ with index set I by letting $x_u = 0$, $x_v = \pi$, and $x_w = -3$. This illustrates several important ways in which nets can differ from sequences.

(a) The index set for a net can be finite.

(b) Nets can have last terms.

(c) Nets can have multiple "first" terms, that is, terms not preceded by other terms.

(d) The index set for a net need not be a chain.

2.1.13 Example. Here is a type of net that is useful in many topological arguments. Suppose that X is a topological space and that $x \in X$. Let I be the collection of all neighborhoods of x with the relation $\preceq$ given by declaring that $U \preceq V$ when $U \supseteq V$. Then I is a directed set. If $x_U \in U$ for each U in I, then (x_U) is a net in X.

A sequence in a topological space converges to an element of the space if, for every neighborhood of that element, all terms of the sequence from some term onward lie in that neighborhood. This definition generalizes immediately to nets.

2.1.14 Definition. Let $(x_\alpha)_{\alpha \in I}$ be a net in a topological space X and let x be an element of X. Then (x_α) *converges* to x, and x is called a *limit* of (x_α), if, for each neighborhood U of x, there is an α_U in I such that $x_\alpha \in U$ whenever $\alpha_U \preceq \alpha$. This convergence is denoted by writing $x_\alpha \to x$ or $\lim_\alpha x_\alpha = x$.[4]

Thus, the net in Example 2.1.13 converges to x. As another example, every bounded increasing function from $\mathbb{R}$ into $\mathbb{R}$, viewed as a net in the sense of Example 2.1.10, converges to its supremum. Notice that the net in Example 2.1.12 converges to -3.

Only subbasic neighborhoods really need to be checked for the property required of all neighborhoods of the point x in Definition 2.1.14.

2.1.15 Proposition. *Suppose that $\mathfrak{S}$ is a subbasis for the topology of a topological space X, that $(x_\alpha)_{\alpha \in I}$ is a net in X, and that $x \in X$. Then*

[4] See footnote 2 on page 141.

$x_\alpha \to x$ if and only if the following is true: For every member U of $\mathfrak{S}$ that contains x, there is an α_U in I such that $x_\alpha \in U$ whenever $\alpha_U \preceq \alpha$.

PROOF. The forward implication follows immediately from the definition of net convergence. For the converse, suppose that every member U of $\mathfrak{S}$ that contains x satisfies the stated condition. It follows from conditions (2) and (3) in Definition 2.1.8 that every finite subset of I has an upper bound in I, which implies that if $\mathfrak{F}$ is a finite subset of $\mathfrak{S}$ each of whose members contains x, then there is an $\alpha_{\mathfrak{F}}$ in I such that $x_\alpha \in \bigcap\{ U : U \in \mathfrak{F} \}$ whenever $\alpha_{\mathfrak{F}} \preceq \alpha$. Therefore $x_\alpha \to x$ because of the way that $\mathfrak{S}$ determines a basis for the topology of X. ■

The proof of the preceding result illustrates a typical application of a useful consequence of conditions (2) and (3) in the definition of a directed set: If (x_α) is a net and $\{P_1, \dots, P_n\}$ is a finite collection of properties defined for the terms of the net such that each P_j holds from some corresponding net index value α_j onward, then there is an index value α such that $P_1, \dots, P_n$ all hold from α onward.

For nets in a topological product, convergence is equivalent to coordinatewise convergence.

2.1.16 Proposition. *Let $\{ X^{(\alpha)} : \alpha \in I \}$ be a family of topological spaces and let X be their topological product. Suppose that $(x_\beta)_{\beta \in J}$ is a net in X and x is a member of X. Then $x_\beta \to x$ if and only if $x_\beta^{(\alpha)} \to x^{(\alpha)}$ for each α in I.*

PROOF. Let $\mathfrak{S}$ be the usual subbasis for the topology of X, that is, the collection of all subsets of X of the form $\prod_{\alpha \in I} U^{(\alpha)}$ such that each $U^{(\alpha)}$ is open and at most one is not equal to the corresponding $X^{(\alpha)}$. It follows from Proposition 2.1.15 that $x_\beta \to x$ if and only if the following holds: For every member U of $\mathfrak{S}$ that contains x, there is a β_U in J such that $x_\beta \in U$ whenever $\beta_U \preceq \beta$. This last condition is equivalent to requiring that $x_\beta^{(\alpha)} \to x^{(\alpha)}$ whenever $\alpha \in I$. ■

Recall that in a Hausdorff space, convergent sequences have unique limits. The corresponding statement for nets actually characterizes Hausdorff spaces among all topological spaces.

2.1.17 Proposition. *A topological space X is a Hausdorff space if and only if each convergent net in X has only one limit.*

PROOF. Suppose that a net (x_α) in X has two different limits x and y. If U_x and U_y are neighborhoods of x and y respectively, then the entire net from some term onward lies in both U_x and U_y, so $U_x \cap U_y \neq \varnothing$. Thus, the space X is not Hausdorff.

Conversely, suppose that X is not Hausdorff. Let x and y be distinct elements of X that cannot be separated by open sets. If U_1 and U_2 are neighborhoods of x and V_1 and V_2 are neighborhoods of y such that $U_1 \supseteq U_2$ and $V_1 \supseteq V_2$, then declare that $(U_1, V_1) \preceq (U_2, V_2)$. For each neighborhood U of x and each neighborhood V of y let $x_{(U,V)}$ be an element of $U \cap V$. Then the net $(x_{(U,V)})$ converges to both x and y. ■

In a metric space, a point is in the closure of a set if and only if some sequence from the set converges to that point. If sequences are replaced by nets, then this remains true for arbitrary topological spaces.

2.1.18 Proposition. *Let S be a subset of a topological space X and let x be an element of X. Then $x \in \overline{S}$ if and only if some net in S converges to x.*

PROOF. If a net in S converges to a point x, then every neighborhood of x includes part of the net and thus intersects S, so $x \in \overline{S}$. Conversely, suppose that $x \in \overline{S}$. Let I be the collection of all neighborhoods of x directed by declaring that $U \preceq V$ when $U \supseteq V$. For each U in I, let x_U be a member of $U \cap S$. Then (x_U) is a net in S converging to x. ■

Since a point x is a limit point of a set S if and only if $x \in \overline{S \setminus \{x\}}$, this next corollary follows immediately from the proposition.

2.1.19 Corollary. *Let S be a subset of a topological space X. Then an element x of X is a limit point of S if and only if there is a net in $S \setminus \{x\}$ converging to x.*

A set in a topological space is closed if and only if it includes its closure. Combining this with the preceding proposition gives the following result. Notice that it generalizes the fact that sets in metric spaces are closed exactly when they are sequentially closed.

2.1.20 Proposition. *A subset S of a topological space is closed if and only if S contains every limit of every net whose terms lie in S.*

The next result is a generalization of the equivalence of continuity and sequential continuity for functions from a metric space into a topological space.

2.1.21 Proposition. *Let X and Y be topological spaces and let f be a function from X into Y.*

(a) *The function f is continuous at the point x_0 of X if and only if $f(x_\alpha) \to f(x_0)$ whenever (x_α) is a net in X converging to x_0.*

(b) *The function f is continuous on X if and only if $f(x_\alpha) \to f(x)$ whenever (x_α) is a net in X converging to an x in X.*

PROOF. Fix an x_0 in X. If f is continuous at x_0 and (x_α) is a net in X converging to x_0, then it follows immediately from the definitions of net convergence and continuity at a point that $f(x_\alpha) \to f(x_0)$. Conversely, suppose that f is not continuous at x_0. Let V be a neighborhood of $f(x_0)$ such that no neighborhood U of x_0 has the property that $f(U) \subseteq V$. Let I be the collection of all neighborhoods of x_0 directed by declaring that $U_1 \preceq U_2$ when $U_1 \supseteq U_2$. For each U in I, let x_U be an element of U such that $f(x_U) \notin V$. Then the net (x_U) converges to x_0, but $\big(f(x_U)\big)$ does not converge to $f(x_0)$. This proves (a), from which (b) follows immediately. ■

2.1.22 Corollary. *If two topologies on the same set result in the same convergent nets with the same limits for those nets, then the two topologies are the same.*

PROOF. Under the hypotheses of the corollary, the identity map on the space, treated as a map between the two topological spaces in question, is continuous in each direction and so is a homeomorphism. ■

2.1.23 Corollary. *Suppose that X, Y, and Z are topological spaces and that $(x,y) \mapsto x \cdot y$ is a continuous operation from $X \times Y$ into Z. Suppose further that (x_α) and (y_α) are nets in X and Y respectively having the same index set, and that x and y are elements of X and Y respectively such that $x_\alpha \to x$ and $y_\alpha \to y$. Then $x_\alpha \cdot y_\alpha \to x \cdot y$.*

PROOF. By Proposition 2.1.16, the net $\big((x_\alpha, y_\alpha)\big)$ in $X \times Y$ converges to (x,y), so an application of Proposition 2.1.21 (b) finishes the proof. ■

A subset of a metric space is compact if and only if it is sequentially compact. In order to generalize this fact to arbitrary topological spaces, some notion for a subnet of a net is needed. By analogy with sequences, it might seem reasonable to let a subnet be the object obtained after carefully thinning the index set of a net. The following definition is needed before exploring this possibility.

2.1.24 Definition. A subset J of a directed set I is *cofinal* in I if for each α in I there is a β_α in J such that $\alpha \preceq \beta_\alpha$.

Suppose that J is a cofinal subset of a directed set I. Then each two elements of J have a common upper bound in I, which by the cofinality of J in I gives them a common upper bound in J. It follows that J is itself a directed set. Now suppose that I is the index set of a net (x_α) in a topological space and that (x_α) has a limit x. Let (x_β) be the net formed by restricting (x_α) to J. A moment's thought about the definitions of net convergence and cofinality shows that (x_β) also converges to x. The fact that J is cofinal in I is important for this. If J were only a subset of I known to be a directed set under the inherited preorder, then (x_β) might have no limit at all.

2.1.25 Example. Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, directed by letting $\mathbb{N}$ retain its natural order and declaring that $\alpha \preceq 0$ for each α in $\mathbb{N}_0$. Define a net from $\mathbb{N}_0$ into $\mathbb{R}$ by letting $x_\alpha = \alpha$ for each α in $\mathbb{N}_0$. Then $x_\alpha \to 0$ though $(x_n)_{n\in\mathbb{N}}$ is a net that does not converge to anything. Notice that $\mathbb{N}$ is not cofinal in $\mathbb{N}_0$.

Let $(x_\alpha)_{\alpha\in I}$ be a net in a topological space. Whatever meaning is assigned to the term *subnet,* any limit that (x_α) has should also be a limit for each of its subnets. Sequential intuition, tempered by Example 2.1.25 and the discussion preceding it, suggests that a subnet of (x_α) should be defined to be the restriction of the net to a cofinal subset J of I, where J has the preorder that it inherits from I. It would seem that it might be necessary to impose additional restrictions on J, but there is no need to worry about this because *this attempt at a definition is already in serious trouble.* As it stands, this preliminary definition has already forced subnets of sequences to be subsequences. This will not do if the net analog of sequential compactness is to be equivalent to compactness, for there are compact Hausdorff spaces in which sequences need not have convergent subsequences; see Exercise 2.14. If such sequences, treated as nets, are to have convergent subnets, then the subnets must be more general objects than subsequences.

The basic problem with this failed attempt to define subnets is the insistence on retaining the full strength of the preorder of I inherited by its cofinal subset J. This is not at all necessary. Suppose instead that J is given a preorder $\preceq_J$ that makes it into a directed set, but not necessarily the preorder $\preceq_I$ inherited from I, and that the net (x_β) formed by restricting (x_α) to J and using the preorder of J is to be a "subnet" of (x_α). If (x_β) is to inherit whatever limits (x_α) has, then it is fairly clear that one should impose on J the requirement that $\beta_1 \preceq_J \beta_2 \Rightarrow \beta_1 \preceq_I \beta_2$. However, there is no particular reason to inflict the converse of that requirement on J. Suppose that $\preceq_J$ is obtained by starting with $\preceq_I$, then weakening that relation on J by declaring that some pairs of elements of J related in I are unrelated in J, but leaving the relation strong enough that J remains a directed set. It is easy to see that the "subnet" (x_β) indexed by J with its weakened relation still has every limit that the full net (x_α) has. In fact, if $\beta_1, \beta_2 \in J$ and $\beta_1 = \beta_2$ in I, then in some sense there is not even any reason to require that $\beta_1 = \beta_2$ in J or even that $\beta_1 \preceq_J \beta_2$ or $\beta_2 \preceq_J \beta_1$ in J! This is the motivation behind the following definition.

2.1.26 Definition. Suppose that X is a set, that I is a directed set, and that $f\colon I \to X$ is a net. Suppose furthermore that J is a directed set and that $g\colon J \to I$ is a function such that

(1) $g(\beta_1) \preceq g(\beta_2)$ in I whenever $\beta_1 \preceq \beta_2$ in J;

(2) $g(J)$ is cofinal in I.

Then the net $f \circ g : J \to X$ is called a *subnet* of f.

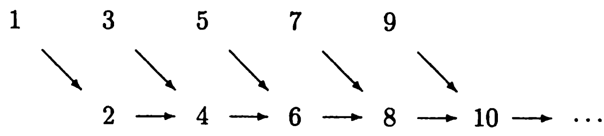


FIGURE 2.1. A weakening of the natural order of $\mathbb{N}$.

The function g in the preceding definition can be treated as a way of embedding J into I as a cofinal subset so that both the relation $\preceq$ and the very notion of equality can be weaker in J itself than in J viewed as a subset of I. If the net f is denoted by (x_α) in the usual way, then its subnet $f \circ g$ should logically be denoted by $(x_{g(\beta)})$. Though clarity sometimes requires that notation, the subnet is often denoted by just (x_β) when no confusion can result. Notice that the original naive attempt to define subnets following Example 2.1.25 does in fact always yield subnets; it just does not yield *all* subnets. In particular, it still follows that a subsequence of a sequence is a subnet of that net. As the next example shows, sequences also have subnets that are not subsequences.

2.1.27 Example. Let (x_n) be any sequence and let J be the natural numbers with their usual order except that each relation between two integers that requires an integer to be properly less than an odd integer has been discarded. The directed set J is represented visually in Figure 2.1, where $\alpha \to \beta$ means that $\alpha \preceq \beta$. If g is the "identity" map from J into the natural numbers with their usual order, then $(x_{g(n)})$ is a subnet of (x_n). Thus, subnets of sequences do not have to be sequences, or even have chains for index sets.

2.1.28 Example. Suppose that (x_n) is any sequence. Let the subset $[1, \infty)$ of $\mathbb{R}$ have its natural order and let $g\colon [1, \infty) \to \mathbb{N}$ be the function that maps each r in $[1, \infty)$ to the greatest integer less than or equal to r. Then (x_r) (that is, $(x_{g(r)})$) is a subnet of (x_n). While a subsequence is often thought of as being obtained from a sequence by thinning the index set, this example shows that a subnet can be formed from a net by thickening the index set (though of course the *range* of a subnet is never any larger than that of the original net). Incidentally, notice that in this particular case the net (x_n) can also be viewed as a subnet of (x_r) in an obvious way.

2.1.29 Example. Let (x_α) be the net of Example 2.1.25. Let J be a one-element set $\{\beta\}$ made into a directed set in the only way possible, and define $g\colon J \to \mathbb{N}_0$ by letting $g(\beta) = 0$. Then $(x_{g(\beta)})$ is a subnet of (x_α). More generally, suppose that J is a directed set and $g\colon J \to \mathbb{N}_0$ is a function to be used to construct a subnet of (x_α). Then the range of g must contain 0 but does not have to contain anything else.

2.1.30 Example. Let f be a continuous real-valued function on $\mathbb{C}$. Let I be the set of all positive reals directed by declaring that $s \preceq t$ when $s \geq t$; that is, the relation $\preceq$ reverses the usual order of the positive reals. For each r in I, let U_r be the open disk in $\mathbb{C}$ of radius r centered at 0 and let S_r be the supremum of f on U_r. Then (S_r) is a convergent net with limit $f(0)$. Let J be the collection of all neighborhoods of 0 in $\mathbb{C}$ besides $\mathbb{C}$ itself, directed by declaring that $U \preceq V$ when $U \supseteq V$. Define $g \colon J \to I$ by letting $g(U) = \max\{ r : U \supseteq U_r \}$. Then $(S_{g(U)})$ is a subnet of (S_r) and converges to $f(0)$.

Here is a collection of facts about subnets and the ways in which their convergence is related to that of the main net.

2.1.31 Proposition. *Let (x_α) be a net in a set X.*

(a) *The net (x_α) is a subnet of itself.*

(b) *Every subnet of (x_α) is a net in X.*

(c) *Every subnet of a subnet of (x_α) is a subnet of (x_α).*

(d) *If X is a topological space and (x_α) converges to an element x of X, then every subnet of (x_α) converges to x.*

(e) *If X is a topological space and there is an element x of X such that every subnet of (x_α) has a subnet converging to x, then $x_\alpha \to x$.*

PROOF. Parts (a), (b), (c), and (d) follow easily from the appropriate definitions. For (e), suppose that X is a topological space and that x is an element of X that is not a limit of (x_α). Then there is a neighborhood U of x with this property: For every α in the index set I for (x_α) there is a β_α in I such that $\alpha \preceq \beta_\alpha$ and $x_{\beta_\alpha} \notin U$. Let $J = \{ \beta : \beta \in I,\, x_\beta \notin U \}$, a cofinal subset of I, and let (x_β) be the restriction of (x_α) to J. Then (x_β) is a subnet of (x_α) that clearly has no subnet converging to x, which proves (e). ∎

2.1.32 Technique. Suppose that (x_α) and (y_β) are nets with respective index sets I and J. It is often useful to be able to find subnets (x_γ) and (y_γ) of (x_α) and (y_β) respectively that have the same index set K. To do this, let $K = I \times J$, directed by declaring that $(\alpha_1, \beta_1) \preceq (\alpha_2, \beta_2)$ when $\alpha_1 \preceq \alpha_2$ and $\beta_1 \preceq \beta_2$. Let $g \colon K \to I$ and $h \colon K \to J$ be the projection mappings, that is, the mappings defined by the formulas $g(\alpha, \beta) = \alpha$ and $h(\alpha, \beta) = \beta$. Then $(x_{g(\alpha,\beta)})$ and $(y_{h(\alpha,\beta)})$ are subnets of (x_α) and (y_β) respectively having the same index set. Notice that these subnets are, in a sense, formed by thickening the index sets of the corresponding nets. Notice also that if (x_α) lies in a topological space, then $(x_{g(\alpha,\beta)})$ converges to some x if *and only if* (x_α) converges to x, and similarly for $(y_{h(\alpha,\beta)})$ and (y_β).

One more notion for sequences needs to be extended to nets before proceeding. Recall that an accumulation point of a sequence in a topological space is a point x such that every neighborhood of x contains terms of the sequence with arbitrarily large indices. The following is the natural generalization to nets.

2.1.33 Definition. Let $(x_\alpha)_{\alpha \in I}$ be a net in a topological space X and let x be an element of X. Then (x_α) *accumulates* at x, and x is called an *accumulation point* of (x_α), if, for each neighborhood U of x and each α in I, there is a $\beta_{\alpha,U}$ in I such that $\alpha \preceq \beta_{\alpha,U}$ and $x_{\beta_{\alpha,U}} \in U$.

Two of the most basic facts about net accumulation are contained in the next result. The proof follows very easily from the relevant definitions.

2.1.34 Proposition. *Suppose that (x_α) is a net in a topological space X and that $x \in X$.*

(a) *If (x_α) converges to x, then (x_α) accumulates at x.*
(b) *If (x_α) has a subnet that accumulates at x, then (x_α) accumulates at x.*

Thus, while convergence is a property passed down from nets to subnets, the property of accumulation is passed up to nets from subnets.

In a metric space, a sequence accumulates at a point if and only if the sequence has a subsequence converging to that point, from which it follows that a set in a metric space is closed exactly when it contains every accumulation point of every sequence in the set. Compare those facts to the following proposition and corollary.

2.1.35 Proposition. *A net in a topological space accumulates at a point if and only if the net has a subnet converging to that point.*

PROOF. Let $(x_\alpha)_{\alpha \in I}$ be a net in a topological space. If (x_α) has a subnet converging to a point x, then that subnet accumulates at x, so (x_α) accumulates at x. For the converse, suppose that (x_α) accumulates at x. Let J be the collection of all ordered pairs (α, U) such that $\alpha \in I$ and U is a neighborhood of x containing x_α. Define a relation on J by declaring that $(\alpha_1, U_1) \preceq (\alpha_2, U_2)$ when $\alpha_1 \preceq \alpha_2$ and $U_1 \supseteq U_2$. If $(\alpha_1, U_1), (\alpha_2, U_2) \in J$, then the fact that (x_α) accumulates at x assures that there is an α_3 such that $\alpha_1 \preceq \alpha_3$, $\alpha_2 \preceq \alpha_3$, and $x_{\alpha_3} \in U_1 \cap U_2$, which implies that $(\alpha_1, U_1) \preceq (\alpha_3, U_1 \cap U_2)$ and $(\alpha_2, U_2) \preceq (\alpha_3, U_1 \cap U_2)$. It follows that this relation defined on J makes J into a directed set. Let $g(\alpha, U) = \alpha$ whenever $(\alpha, U) \in J$. Then $(x_{g(\alpha,U)})$ is a subnet of (x_α) converging to x. ■

2.1.36 Corollary. *A subset S of a topological space is closed if and only if S contains every accumulation point of every net whose terms lie in S.*

PROOF. It follows from Proposition 2.1.20 that S is closed if and only if it contains every limit of every convergent subnet of every net whose terms lie in S, so an application of the proposition just proved yields this corollary. ■

The following result is the analog for arbitrary topological spaces of the equivalence of compactness and sequential compactness in metric spaces.

2.1.37 Proposition. *A subset S of a topological space is compact if and only if each net in S has a subnet with a limit in S, that is, if and only if each net in S has an accumulation point in S.*

PROOF. Suppose that (x_α) is a net in S with no accumulation point in S. For each x in S, let U_x be a neighborhood of x that excludes the entire portion of the net from some term onward. Let $\mathfrak{G} = \{ U_x : x \in S \}$, an open covering for S. Since every finite subcollection of $\mathfrak{G}$ excludes the entire net from some term onward, it follows that $\mathfrak{G}$ cannot be thinned to a finite subcovering for S, so S is not compact.

Conversely, suppose that S has an open covering $\mathfrak{G}$ that cannot be thinned to a finite subcovering for S. It can be assumed that $\mathfrak{G}$ is closed under the operation of taking finite unions of its elements. It follows that $\mathfrak{G}$ can be made into a directed set by declaring that $U \preceq V$ when $U \subseteq V$. For each U in $\mathfrak{G}$, let x_U be a member of $S \setminus U$. Then (x_U) is a net in S with the property that $x_{U_2} \notin U_1$ when $U_1 \preceq U_2$. It follows that (x_U) has no accumulation point in S. ■

Compactness also has a useful characterization in terms of the convergence of special nets called *ultranets*. See Appendix D for this characterization as well as a general discussion of ultranets and examples of how they can be used to simplify compactness arguments that involve the axiom of choice.

By analogy with metric spaces, it would seem reasonable to conjecture that a subset S of a topological space X must be relatively compact if each net in S has a subnet with a limit in X. Proposition 2.1.37 would seem to provide strong supporting evidence for this conjecture. It is perhaps surprising that the conjecture is in general false; see Exercise 2.15. It does hold, however, for a class of topological spaces large enough to include all of the important ones that appear in this book.

2.1.38 Definition. Suppose that X is a set with a *group (multiplication) operation*, that is, an operation $(x, y) \mapsto x \cdot y$ from $X \times X$ into X such that

(1) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ whenever $x, y, z \in X$; that is, the operation is associative;

(2) there is an *identity element* e in X such that $x \cdot e = e \cdot x = x$ whenever $x \in X$;

(3) each element x of X has an *inverse* x^{-1} in X such that $x \cdot x^{-1} = x^{-1} \cdot x = e$.

Then $(X,\cdot)$ is a *group*. Suppose furthermore that $\mathfrak{T}$ is a topology for X such that the mappings $(x,y) \mapsto x \cdot y$ from $X \times X$ into X and $x \mapsto x^{-1}$ from X into X are both continuous. Then $(X,\mathfrak{T},\cdot)$ is a *topological group.* When no confusion can result, both the group $(X,\cdot)$ and the topological group $(X,\mathfrak{T},\cdot)$ are denoted by X.

For example, if X is a normed space and the group operation is addition of vectors, then X is a topological group by Proposition 1.3.2.

The reader should be warned that other sources sometimes define a topological group to be a Hausdorff space satisfying Definition 2.1.38. Most of the topological groups important in analysis actually are Hausdorff spaces. In fact, a topological group that is not a Hausdorff space is not even a T_0 space; see Exercise 2.16.

Recall that a group is said to be *abelian* if the group operation is commutative. When this happens, the operation is often called *group addition* instead of multiplication, with a corresponding shift to additive notation such as $x+y$ and $-x$. In particular, the identity of an abelian group is often denoted by 0 instead of e. Most of the topological groups encountered in this book will be abelian.

A bit more notation is needed before proceeding. Suppose that X is a group and that $x \in X$ and $A, B \subseteq X$. Then $x \cdot A = \{x \cdot a : a \in A\}$, $A \cdot x = \{a \cdot x : a \in A\}$, $A \cdot B = \{a \cdot b : a \in A, b \in B\}$, and $A^{-1} = \{a^{-1} : a \in A\}$. When the group is abelian and additive notation is being used, the corresponding objects are denoted by $x+A$, $A+x$, $A+B$, and $-A$.

2.1.39 Proposition. *Suppose that X is a topological group with identity e.*

(a) *Let x_0 be an element of X. Then the maps $x \mapsto x_0 \cdot x$, $x \mapsto x \cdot x_0$, and $x \mapsto x^{-1}$ are homeomorphisms from X onto itself. Consequently, if A is a subset of X that is open, or closed, or compact, then $x_0 \cdot A$, $A \cdot x_0$, and A^{-1} also have that property. If A and G are subsets of X and G is open, then both $A \cdot G$ and $G \cdot A$ are open.*

(b) *For each x_0 in X, the neighborhoods of x_0 are exactly the sets $x_0 \cdot U$ such that U is a neighborhood of e, which are in turn exactly the sets $U \cdot x_0$ such that U is a neighborhood of e.*

(c) *For each neighborhood U of e, there is a neighborhood V of e such that $V = V^{-1}$ and $V \cdot V \subseteq U$.*

PROOF. For (a), suppose that $x_0 \in X$. The continuity of the group operations of X implies that the maps $x \mapsto x_0 \cdot x$, $x \mapsto x \cdot x_0$, and $x \mapsto x^{-1}$ are continuous, as are their respective inverses $x \mapsto x_0^{-1} \cdot x$, $x \mapsto x \cdot x_0^{-1}$, and $x \mapsto x^{-1}$, so the three maps of part (a) are homeomorphisms from X onto itself. If A is a subset of X that is open, or closed, or compact, then $x_0 \cdot A$, $A \cdot x_0$, and A^{-1} have that same property since these properties are preserved by homeomorphisms. If A and G are subsets of X and G is open, then $A \cdot G$

and $G \cdot A$ are open as unions of open sets, since $A \cdot G = \bigcup\{ a \cdot G : a \in A \}$ and $G \cdot A = \bigcup\{ G \cdot a : a \in A \}$. This proves (a). If U_e and U_{x_0} are neighborhoods of e and x_0 respectively, then (a) implies that $x_0 \cdot U_e$ and $x_0^{-1} \cdot U_{x_0}$ are neighborhoods of x_0 and e respectively, which together with the fact that $U_{x_0} = x_0 \cdot (x_0^{-1} \cdot U_{x_0})$ easily yields the first part of (b). The second part is proved similarly.

Finally, suppose that U is a neighborhood of e. By the continuity of group multiplication, there are neighborhoods V_1 and V_2 of e such that $V_1 \cdot V_2 \subseteq U$. Let $V = V_1 \cap V_2 \cap V_1^{-1} \cap V_2^{-1}$, another neighborhood of e. Then V does what is required of it in (c). ■

Here is the analog of Proposition 2.1.37 for relative compactness in topological groups that could not be obtained for arbitrary topological spaces.

2.1.40 Proposition. *A subset S of a topological group X is relatively compact if and only if each net in S has a subnet with a limit in X (not assumed to be in S), that is, if and only if each net in S has an accumulation point in X.*

PROOF. If S is relatively compact, then $\overline{S}$ is compact, so every net in S has a subnet with a limit in $\overline{S}$ and therefore in X.

For the converse, suppose that every net in S has a subnet with a limit in X. Let $(x_\alpha)_{\alpha \in I}$ be a net in $\overline{S}$. By Propositions 2.1.20 and 2.1.37, it is enough to show that (x_α) has a convergent subnet. For each α in I and each neighborhood U of the identity e of X, let $y_{(\alpha,U)}$ be an element of $(x_\alpha \cdot U) \cap S$. Then $(y_{(\alpha,U)})$ is a net if its index set J is preordered by declaring that $(\alpha_1, U_1) \preceq (\alpha_2, U_2)$ when $\alpha_1 \preceq \alpha_2$ and $U_1 \supseteq U_2$. Furthermore, a corresponding subnet $(x_{(\alpha,U)})$ of (x_α) is obtained by letting $x_{(\alpha,U)} = x_\alpha$ for each (α, U) in J. It is enough to show that $(x_{(\alpha,U)})$ has a convergent subnet. Since the net $(y_{(\alpha,U)})$ has a subnet (y_β) with a limit x, it is enough to show that the corresponding subnet (x_β) of $(x_{(\alpha,U)})$ also converges to x.

Suppose that U_0 is a neighborhood of e. Let α_0 be any element of I. If $(\alpha_0, U_0) \preceq (\alpha, U)$, then $y_{(\alpha,U)} \in x_{(\alpha,U)} \cdot U_0$, so $x_{(\alpha,U)}^{-1} \cdot y_{(\alpha,U)} \in U_0$. It follows that $x_{(\alpha,U)}^{-1} \cdot y_{(\alpha,U)} \to e$ and therefore that $x_\beta^{-1} \cdot y_\beta \to e$. The continuity of the group operations then assures that

$$y_\beta^{-1} \cdot x_\beta = (x_\beta^{-1} \cdot y_\beta)^{-1} \to e^{-1} = e$$

and that

$$x_\beta = y_\beta \cdot (y_\beta^{-1} \cdot x_\beta) \to x \cdot e = x,$$

as required. ■

See Proposition D.10 in Appendix D for a characterization of relative compactness in topological groups in terms of the convergence of ultranets.

In metric spaces, the notion of a Cauchy sequence has an obvious generalization to nets: A net $(x_\alpha)_{\alpha\in I}$ in a metric space is *Cauchy* if, for every positive ϵ, there is an α_ϵ in I such that $d(x_\beta, x_\gamma) < \epsilon$ whenever $\alpha_\epsilon \preceq \beta$ and $\alpha_\epsilon \preceq \gamma$. Because of the special role played by the metric in this definition, it is not quite so easy to obtain a further generalization to more abstract topological settings. It turns out to be more fruitful to start with the special case of the Cauchy condition for nets in the real line, concentrate on the topological aspects of that condition rather than the metric space aspects, and then attempt a generalization from that special case.

Suppose that $(r_\alpha)_{\alpha\in I}$ is a net in $\mathbb{R}$. Then $\{r_\beta - r_\gamma : (\beta,\gamma) \in I \times I\}$ becomes a net if $I \times I$ is preordered by declaring that $(\beta_1,\gamma_1) \preceq (\beta_2,\gamma_2)$ when $\beta_1 \preceq \beta_2$ and $\gamma_1 \preceq \gamma_2$. The Cauchy condition for nets then becomes a convergence statement about such a net of differences: The net (r_α) is Cauchy if and only if the difference net $(r_\beta - r_\gamma)_{(\beta,\gamma)\in I\times I}$ converges to 0.

This topological characterization of the Cauchy condition for nets in $\mathbb{R}$ can be immediately generalized to an abelian group X with a topology by declaring that a net $(x_\alpha)_{\alpha\in I}$ in X is Cauchy if, under the same indexing scheme as for difference nets in $\mathbb{R}$, the difference net $(x_\beta - x_\gamma)_{(\beta,\gamma)\in I\times I}$ converges to the identity element 0 of the group. To have a useful theory of Cauchy nets in such a topological space, it is reasonable to ask that convergent nets be Cauchy. By analogy with the usual argument for sequences in the real line, the proof that a net $(x_\alpha)_{\alpha\in I}$ in X with a limit x_0 is Cauchy should amount to showing that for every neighborhood U of 0 there is an α_U in I such that if $\alpha_U \preceq \beta$ and $\alpha_U \preceq \gamma$, then x_β is "near enough" to x_0 and $-x_\gamma$ is "near enough" to $-x_0$ to force $x_\beta - x_\gamma$ to be "near enough" to the sum 0 of x_0 and $-x_0$ to lie in U. (All of this language will be made rigorous in the proof of Proposition 2.1.47.) To assure that this will happen, the maps $(x, y) \mapsto x + y$ from $X \times X$ into X and $x \mapsto -x$ from X into X should be continuous; that is, the space X should be a topological group.

Thus, abelian topological groups form a natural setting for a generalization of the Cauchy condition. To avoid direct references to the convergence of difference nets, this generalization can be stated as follows.

2.1.41 Definition. A net $(x_\alpha)_{\alpha\in I}$ in an abelian topological group X is *(topologically) Cauchy* if, for every neighborhood U of the identity element 0 of X, there is an α_U in I such that $x_\beta - x_\gamma \in U$ whenever $\alpha_U \preceq \beta$ and $\alpha_U \preceq \gamma$.

It is worth emphasizing that the preceding definition is *not* a generalization of the Cauchy condition for nets in metric spaces. It is instead a generalization of the topological formulation of the Cauchy condition for nets in the real line, while the Cauchy condition for nets in metric spaces is a different generalization based on the metric formulation of the Cauchy condition in $\mathbb{R}$. This has unfortunate consequences when an abelian topological group has its topology induced by a metric, for it is quite possible for a net in such a space to be metrically Cauchy or topologically Cauchy without being Cauchy in both senses; see Exercise 2.17. However, this difficulty does not arise in the situations of most importance in this book, as will be shown by Proposition 2.1.44 and its corollary.

2.1.42 Definition. A metric d on a group X is *left-invariant* (respectively, *right-invariant*) if $d(z\cdot x, z\cdot y) = d(x, y)$ (respectively, $d(x\cdot z, y\cdot z) = d(x, y)$) whenever $x, y, z \in X$. If d is both left-invariant and right-invariant, then d is *invariant.*

It follows that a metric d on an abelian group X is invariant if and only if $d(x + z, y + z) = d(x, y)$ whenever $x, y, z \in X$. In particular, a metric on a vector space is said to be invariant when it has this property, since the space is an abelian group under vector addition. Invariant metrics on abelian groups are sometimes said to be *translation-invariant* because of the additive notation.

This is an appropriate place for the following result, which will be used in several of the examples in the next section.

2.1.43 Proposition. *If a topology for a group is induced by an invariant metric, then the group is a topological group when given this topology.*

PROOF. Suppose that a group X is given a topology that is induced by an invariant metric d. Let e be the identity element of X. If sequences (x_n) and (y_n) converge to x and y respectively in X, then

$$\begin{aligned} d(x_n \cdot y_n, x \cdot y) &= d(x^{-1}\cdot x_n, y \cdot y_n^{-1}) \\ &\le d(x^{-1}\cdot x_n, e) + d(e, y \cdot y_n^{-1}) \\ &= d(x_n, x) + d(y_n, y) \to 0 \end{aligned}$$

and

$$d(x_n^{-1}, x^{-1}) = d(x, x_n) \to 0,$$

so $x_n \cdot y_n \to x \cdot y$ and $x_n^{-1} \to x^{-1}$. It follows that group multiplication and inversion are both continuous, so X is a topological group. ■

2.1.44 Proposition. *Suppose that the topology of an abelian topological group is induced by an invariant metric. Then a net in the group is Cauchy with respect to this metric if and only if the net is topologically Cauchy.*

PROOF. Let (x_α) be a net in an abelian topological group whose topology is induced by an invariant metric d. If x_β and x_γ are two terms of the net, then $d(x_\beta, x_\gamma) = d(x_\beta - x_\gamma, 0)$, from which it easily follows that the net is metrically Cauchy if and only if it is topologically Cauchy. ■

Since every normed space is an abelian topological group when the group operation is vector addition, and since the metric induced by the norm is invariant for this operation, the following result is an immediate consequence of the proposition just proved.

2.1.45 Corollary. *Let X be a normed space, treated as an abelian topological group under vector addition. Then a net in X is topologically Cauchy if and only if it is norm Cauchy.*

In the real line, the convergence of a subsequence of a Cauchy sequence is enough to force the convergence of the entire sequence to the same limit. This fact generalizes to nets in abelian topological groups.

2.1.46 Proposition. *If a Cauchy net in an abelian topological group has a convergent subnet, then the entire net converges to the limit of the subnet.*

PROOF. Suppose that $(x_\alpha)_{\alpha\in I}$ is a Cauchy net in an abelian topological group X and that (x_α) has a subnet with a limit x_0. Let U be a neighborhood of 0 in X, let V be a neighborhood of 0 such that $V+V\subseteq U$, and let α_0 be a member of I such that $x_\gamma - x_\delta \in V$ whenever $\alpha_0 \preceq \gamma$ and $\alpha_0 \preceq \delta$. Since (x_α) accumulates at x_0, there is an α_U in I such that $\alpha_0 \preceq \alpha_U$ and $x_{\alpha_U} \in x_0 + V$. If $\alpha_U \preceq \alpha$, then

$$x_\alpha = (x_\alpha - x_{\alpha_U}) + x_{\alpha_U} \in V + (x_0 + V) = x_0 + (V+V) \subseteq x_0 + U,$$

which implies that $x_\alpha \to x_0$. ■

The remarks that precede Definition 2.1.41 include a sketch of a proof that convergent nets in abelian topological groups are always Cauchy. It is time to fill in the details of that proof.

2.1.47 Proposition. *Every convergent net in an abelian topological group is Cauchy.*

PROOF. Let X be an abelian topological group and let $(x_\alpha)_{\alpha\in I}$ be a net in X converging to some element x_0 of X. Let U be a neighborhood of 0 in X and let V be a neighborhood of 0 such that $V=-V$ and $V+V\subseteq U$. Let α_U be a member of I such that $x_\alpha \in x_0 + V$ whenever $\alpha_U \preceq \alpha$. If $\alpha_U \preceq \beta$ and $\alpha_U \preceq \gamma$, then

$$x_\beta - x_\gamma \in (x_0+V) - (x_0+V) = V - V = V + V \subseteq U,$$

so (x_α) is Cauchy. ■

2.1.48 Definition. An abelian topological group is *complete* if each Cauchy net in the group converges.

This definition of completeness might seem fundamentally different from the usual one for metric spaces, since that definition requires only that all metrically Cauchy *sequences* converge. However, it turns out that the completeness of a metric space in the usual sense is actually enough to guarantee the convergence of all of its metrically Cauchy nets.

2.1.49 Proposition. *Let (X, d) be a metric space. Then d is a complete metric if and only if each d-Cauchy net converges.*

PROOF. Suppose that d is complete and that $(x_\alpha)_{\alpha \in I}$ is a Cauchy net in X. It is enough to prove that (x_α) converges. Since (x_α) is Cauchy, there is a sequence (α_n) in I such that $\alpha_n \preceq \alpha_{n+1}$ for each n and $d(x_\beta, x_\gamma) < n^{-1}$ whenever $\alpha_n \preceq \beta$ and $\alpha_n \preceq \gamma$. It follows that (x_{α_n}) is a Cauchy sequence and so has a limit x. If $\epsilon > 0$, then there is a positive integer m such that $m > 2/\epsilon$ and $d(x_{\alpha_m}, x) < \epsilon/2$, which implies that if $\alpha_m \preceq \alpha$, then

$$d(x_\alpha, x) \le d(x_\alpha, x_{\alpha_m}) + d(x_{\alpha_m}, x) < \frac{1}{m} + \frac{\epsilon}{2} < \epsilon.$$

The net (x_α) therefore converges to x. ∎

Combining the preceding proposition with Proposition 2.1.44 yields the following result immediately.

2.1.50 Corollary. *Suppose that the topology of an abelian topological group X is induced by an invariant metric d. Then X is complete as an abelian topological group if and only if d is a complete metric.*

2.1.51 Corollary. *Let X be a normed space. Then X is complete as an abelian topological group under vector addition if and only if X is a Banach space.*

In light of Proposition 2.1.49, one might ask if the convergence of every Cauchy sequence in an abelian topological group is enough to assure that the group is complete. As will be seen in Example 2.5.24, the answer is no.

Throughout the rest of this book, nets are the objects whose convergence properties are used in topological arguments. They are not the only possible choice for this. Many authors use the convergence properties of filters or their close relatives, filterbases, to accomplish the same thing. This is especially true for those who learned much of their mathematics from Nicolas Bourbaki's *Éléments de Mathématique* [33] since Bourbaki uses filters to do "his" topology. (Those who do not know why the quotes are around the word "his" should read Paul Halmos's delightful *Scientific American* article [100] about Bourbaki.) Suffice it to say here that whenever a net converges, there are related filters and filterbases that are also forced to converge, and that if a filter or filterbase converges, then there is a convergent net lurking not too far away. See, for example, [65] for a discussion of filters and filterbases and the relationship of their convergence properties to those of nets.

Exercises

2.1 Let S be a subset of a metric space. Prove the following.

(a) The set S is relatively sequentially compact if and only if its closure is sequentially compact.

(b) The set S is relatively limit point compact if and only if its closure is limit point compact.

2.2 Let X be a topological space.

(a) Prove that X is a T_1 space if and only if each one-element subset of X is closed.

(b) Prove that X is regular if and only if it is a T_1 space with this property: Whenever $x \in X$ and U is a neighborhood of x, there is a neighborhood V of x whose closure lies in U.

(c) Prove that X is normal if and only if it is a T_1 space with this property: Whenever F is a closed subset of X and U is an open superset of F, there is an open superset V of F whose closure lies in U.

2.3 Let X be the interval $[0, 1]$ with the topology given by declaring that a subset of X is open if it does not contain 0 or its complement is countable. Verify that X is a Hausdorff topological space. Show that the subset $(0, 1]$ of X is sequentially closed but not closed.

2.4 Let A and B be subsets of topological spaces X and Y respectively. Using net arguments, prove that $\overline{A \times B} = \overline{A} \times \overline{B}$ in $X \times Y$.

2.5 Using net arguments, prove that continuous images of compact sets are always compact.

2.6 Using net arguments, prove that every compact subset of a Hausdorff space is closed.

2.7 Using net arguments, prove that a topological space X is Hausdorff if and only if the *diagonal* $\{ (x, x) : x \in X \}$ is closed in $X \times X$.

2.8 Let X be a set. Using net arguments, prove that no Hausdorff topology for X can be properly weaker (that is, have fewer open sets) than a topology for X with respect to which X is compact.

2.9 (a) Let K be a compact subset of a topological space X. Suppose that $(x_\alpha)_{\alpha \in I}$ is a net in X such that whenever U is an open set that includes K and $\alpha \in I$, there is a $\beta_{\alpha,U}$ in I such that $\alpha \preceq \beta_{\alpha,U}$ and $x_{\beta_{\alpha,U}} \in U$. Prove that some subnet of (x_α) converges to an element of K.

(b) Use (a) to prove that every compact Hausdorff space is normal.

2.10 A subset S of a topological space is a *Lindelöf* set if each open covering of S can be thinned to a countable subcovering. Suppose that the net (x_α) in the Lindelöf set S has no subnet converging to any point of S. Show that there is a sequence (α_n) of net indices for (x_α) such that $\alpha_n \preceq \alpha_{n+1}$ for each n and the sequence (x_{α_n}) has no subnet converging to a point of S. Conclude that every sequentially compact Lindelöf set is compact.

2.11 Recall that a topological space X is *connected* if there is no pair of disjoint nonempty open subsets of X whose union is X.

(a) Prove that a topological space X is connected if and only if it has this property: Whenever S is a nonempty proper subset of X, there is either a net in S converging to a point in the complement of S, or a net in the complement of S converging to a point in S.

(b) Let X and Y be topological spaces such that X is connected. Suppose that there is a continuous function from X onto Y. Use (a) to show that Y is connected.

2.12 Let a and b be real numbers such that $a < b$. If T is a finite subset $\{t_0, \dots, t_n\}$ of $[a, b]$ such that $t_0 = a$, $t_n = b$, and $t_{j-1} < t_j$ when $j \in \{1, \dots, n\}$, and S is a corresponding finite subset $\{s_1, \dots, s_n\}$ of $[a, b]$ such that $s_j \in [t_{j-1}, t_j]$ for each j, then the ordered pair (T, S) is called a *Riemann partition* of $[a, b]$. Define a relation $\preceq$ on the collection I of all Riemann partitions of $[a, b]$ by declaring that $(T_1, S_1) \preceq (T_2, S_2)$ when $T_1 \subseteq T_2$. Let f be a real-valued function on $[a, b]$. For each (T, S) in I, let $x_{(T,S)} = \sum_{j=1}^n f(s_j)(t_j - t_{j-1})$.

(a) Show that I with the relation $\preceq$ is a directed set.

(b) Find necessary and sufficient conditions for the net $(x_{(T,S)})$ to converge. When it converges, identify its limit.

2.13 Let a and b be real numbers such that $a < b$ and let λ be Lebesgue measure on the Lebesgue-measurable subsets of $[a, b]$. Let I be the collection of all partitions of $[a, b]$ into a finite number of disjoint measurable sets. Define a relation $\preceq$ on I in the following way. Suppose that $P_1, P_2 \in I$ and that $P_1 = \{E_1, \dots, E_n\}$ and $P_2 = \{F_1, \dots, F_m\}$. Then $P_1 \preceq P_2$ if each E_j is the union of the members of some subcollection of P_2, that is, if the members of P_2 are formed by subdividing the members of P_1. Now let f be a nonnegative real-valued measurable function on $[a, b]$. If $P \in I$ and $P = \{E_1, \dots, E_n\}$, let $x_P = \sum_{j=1}^n \inf\{f(t) : t \in E_j\}\lambda(E_j)$.

(a) Show that I with the relation $\preceq$ is a directed set.

(b) Find necessary and sufficient conditions for the net (x_P) to converge. When it converges, identify its limit.

2.14 Let Y be a two-element set with the *discrete topology*, that is, the topology in which every subset of the set is open. Let X be the collection of all functions from the closed interval $[0, 1]$ into Y, viewed as the topological product $\prod_{\alpha\in[0,1]} Y_\alpha$ where $Y_\alpha = Y$ for each α. Show that X is a compact Hausdorff space that is not sequentially compact. (Notice that if (f_n) is a sequence in X, then the collection of all subsequences of (f_n) can be put into one-to-one correspondence with $(0, 1]$.)

Examples of sequentially compact Hausdorff spaces that are not compact can be found in [228].

2.15 Let $X = \{ z : z \in \mathbb{C}, |z| \leq 1 \}$ and let $S = \{ z : z \in \mathbb{C}, |z| < 1 \}$. For each z in $X \setminus S$, let an *inner deleted neighborhood* of z be a set of the form $U \cap S$, where U is a neighborhood of z in $\mathbb{C}$ in the usual sense. Define a subset G of X to be open if $G \cap S$ is open in $\mathbb{C}$ in the usual sense and for each z in $G \cap (X \setminus S)$ there is an inner deleted neighborhood of z included in G. Show that this defines a Hausdorff topology on X for which $\overline{S}$ is not sequentially compact, limit point compact, or compact. Show that S is relatively sequentially compact and relatively limit point compact, and that each net in S has a subnet with a limit in X.

2.16 Prove that every T_0 topological group is Hausdorff.

2.17 Show that the formula $d(r_1, r_2) = \left|(r_1 - \sqrt{2})^{-1} - (r_2 - \sqrt{2})^{-1}\right|$ defines a metric on the rationals $\mathbb{Q}$ that induces the usual topology that $\mathbb{Q}$ inherits from $\mathbb{R}$. Use this to give an example of an abelian topological group whose topology is induced by a metric such that some sequence in the group is topologically Cauchy but not metrically Cauchy, and some other sequence in the group is metrically Cauchy but not topologically Cauchy.

2.2 Vector Topologies

Some useful properties of normed spaces proved in Chapter 1, such as the fact that the closed convex hull of a set in a normed space is the closure of its convex hull, actually follow only from the continuity of the vector space operations of such spaces rather than from any other special properties of norms. It turns out that a number of other important properties of these spaces, such as the fact that every continuous linear functional on a subspace can be extended continuously to the entire space, also rely on the fact that every neighborhood of a point in a normed space includes a convex neighborhood of that point, but not on any other special properties of normed spaces beyond the continuity of the vector space operations. This suggests the following generalizations of the notion of a normed space.

2.2.1 Definitions. Suppose that X is a vector space with a topology $\mathfrak{T}$ such that addition of vectors is a continuous operation from $X \times X$ into X and multiplication of vectors by scalars is a continuous operation from $\mathbb{F} \times X$ into X. Then $\mathfrak{T}$ is a *vector* or *linear* topology for X, and the ordered pair $(X, \mathfrak{T})$ is a *topological vector space (TVS)* or a *linear topological space (LTS)*. If $\mathfrak{T}$ has a basis consisting of convex sets, then $\mathfrak{T}$ is a *locally convex* topology and the TVS $(X, \mathfrak{T})$ is a *locally convex space (LCS)*.

When $(X, \mathfrak{T})$ is a topological vector space, it is usually said informally (though admittedly not quite correctly) that X is the TVS, unless the more formal notation is needed for clarity.

The continuity of multiplication of vectors by scalars in a topological vector space implies the continuity of the map $x \mapsto -x$ on the space, which immediately yields the following result. This characterization is sometimes used as the definition of a vector topology.

2.2.2 Proposition. *A topology on a vector space X is a vector topology if and only if both of the following occur: The space X is a topological group under addition of vectors, and multiplication of vectors by scalars is a continuous operation from $\mathbb{F} \times X$ into X.*

The continuity of the vector space operations in a TVS creates a link between the vector space structure and the topology of the space. The additional property possessed by an LCS provides each of its points with a supply of nicely shaped neighborhoods. Almost all of the TVSs important in Banach space theory are T_0 spaces, which in fact implies that they are completely regular, as will be proved in Theorem 2.2.14. For this reason, many sources include some separation axiom, ranging from T_0 through $T_{3\frac{1}{2}}$, as one of the defining properties of a vector topology.

As the following result shows, Chapter 1 is teeming with examples of locally convex spaces.

2.2.3 Theorem. *Every norm topology is a locally convex topology.*

PROOF. The continuity of the vector space operations of a normed space is proved in Proposition 1.3.2, while the existence of a basis consisting of convex sets comes from the fact that every ball in a normed space is convex. ■

Here are some examples of vector topologies that are not in general induced by norms.

2.2.4 Example. Let X be a vector space having a nonzero vector. Then the topology $\{\varnothing, X\}$ is a locally convex topology for X that is not even a T_0 topology.

Some abbreviations will be helpful in preventing the notation in the next example from becoming too unwieldy. Suppose that (Ω, Σ, μ) is a measure space, that f is a scalar-valued function on Ω, and that P is a property that $f(x)$ might satisfy for some members x of Ω. Then $[f$ satisfies $P]$ will be used as an abbreviation for $\{x : x \in \Omega, f(x) \text{ satisfies } P\}$. If $[f$ satisfies $P]$ is a measurable set, then $\mu[f$ satisfies $P]$ will be used to denote $\mu([f$ satisfies $P])$.

2.2.5 Example: $L_0(\Omega, \Sigma, \mu)$. Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω. Let $L_0(\Omega, \Sigma, \mu)$ be the vector space of all μ-measurable scalar-valued functions on Ω, with the usual convention that two functions are identified if they agree almost everywhere. It is a straightforward exercise in measure theory to show that the formula

$$d(f,g) = \inf\{\{1\} \cup \{\epsilon : \epsilon > 0,\ \mu[\,|f-g| > \epsilon\,] < \epsilon\}\}$$

defines a metric on $L_0(\Omega, \Sigma, \mu)$, that a sequence in $L_0(\Omega, \Sigma, \mu)$ is Cauchy with respect to d if and only if it is Cauchy in measure, and that a sequence in $L_0(\Omega, \Sigma, \mu)$ converges to an element f of $L_0(\Omega, \Sigma, \mu)$ with respect to d if and only if the sequence converges to f in measure; see Exercise 2.18. Since sequences that are Cauchy in measure are convergent in measure, the metric d is complete. The *topology of convergence in measure* for $L_0(\Omega, \Sigma, \mu)$ is the topology induced by this metric. *Henceforth, this will be the topology that $L_0(\Omega, \Sigma, \mu)$ is assumed to have whenever it is treated as a topological space.*

It is clear that d is invariant, that is, that $d(f+h, g+h) = d(f,g)$ whenever $f, g, h \in L_0(\Omega, \Sigma, \mu)$. It follows from Proposition 2.1.43 that $L_0(\Omega, \Sigma, \mu)$ is a topological group under the operation of addition of vectors, and in particular that vector addition is continuous. However, it is not in general true that multiplication of vectors by scalars is a continuous operation for $L_0(\Omega, \Sigma, \mu)$; see Exercise 2.18. It does turn out that this operation is continuous when μ is a finite measure, as will now be shown.

Assume for the rest of this example that $\mu(\Omega) < \infty$. Suppose that the sequences (α_n) and (f_n) are in $\mathbb{F}$ and $L_0(\Omega, \Sigma, \mu)$ respectively and have respective limits α and f. Let (b_n) be a sequence of positive numbers decreasing to 0 such that $|\alpha_n - \alpha| \le b_n$ for each n, and let B be a positive upper bound for the sequence $\big(|\alpha_n|\big)$. For each positive ϵ and each positive integer n,

$$[\,|\alpha_n f_n - \alpha f| > \epsilon\,] \subseteq \left[\,|\alpha_n f_n - \alpha_n f| > \frac{\epsilon}{2}\right] \cup \left[\,|\alpha_n f - \alpha f| > \frac{\epsilon}{2}\right],$$

from which it follows that

$$\begin{aligned}
\mu[\,|\alpha_n f_n - \alpha f| > \epsilon\,] &\le \mu\left[\,|\alpha_n f_n - \alpha_n f| > \frac{\epsilon}{2}\right] + \mu\left[\,|\alpha_n f - \alpha f| > \frac{\epsilon}{2}\right] \\
&\le \mu\left[\,B|f_n - f| > \frac{\epsilon}{2}\right] + \mu\left[\,b_n|f| > \frac{\epsilon}{2}\right] \\
&= \mu\left[\,|f_n - f| > \frac{\epsilon}{2B}\right] + \mu\left[\,|f| > \frac{\epsilon}{2b_n}\right].
\end{aligned}$$

Letting n tend to ∞ shows that $\mu[\,|\alpha_n f_n - \alpha f| > \epsilon\,] \to 0$ whenever $\epsilon > 0$, so $\alpha_n f_n$ converges to αf in measure. (Notice that the finiteness of $\mu(\Omega)$ is used here to assure that $\mu[\,|f| > \epsilon/2b_n\,] \to 0$.) Multiplication of vectors by scalars is therefore a continuous operation, so the topology of convergence in measure is a vector topology for $L_0(\Omega, \Sigma, \mu)$ whenever (Ω, Σ, μ) is a finite positive measure space.

Now suppose that Ω is the interval $[0,1]$, that Σ is the σ-algebra of Lebesgue-measurable subsets of Ω, and that λ is Lebesgue measure on

(Ω, Σ). Let $L_0[0,1]$ denote $L_0(\Omega, \Sigma, \lambda)$. Let C be a convex neighborhood of 0 in $L_0[0,1]$ and let ϵ be such that the open ball of radius ϵ centered at 0 lies in C. Suppose that $f \in L_0[0,1]$. Let n be a positive integer such that $1/n < \epsilon/2$. For each integer j such that $1 \le j \le n$, let $\mathbf{I}_{[\frac{j-1}{n}, \frac{j}{n})}$ be the indicator function of the interval $[\frac{j-1}{n}, \frac{j}{n})$ and let the member f_j of $L_0[0,1]$ be defined by the formula $f_j(t) = nf(t)\mathbf{I}_{[\frac{j-1}{n}, \frac{j}{n})}(t)$. Then $\lambda[\,|f_j| > \epsilon/2] \le 1/n < \epsilon/2$ for each f_j, so each f_j is in C. Since $f = \sum_{j=1}^{n} n^{-1} f_j$, the convex combination f of members of C is itself in C. It follows that $C = L_0[0,1]$. By Proposition 2.1.39 (b), every nonempty convex open set in $L_0[0,1]$ is a translate of a convex neighborhood of the origin, so it also follows that the only nonempty convex open subset of $L_0[0,1]$ is $L_0[0,1]$ itself. Thus, the topology of convergence in measure for $L_0[0,1]$ is a vector topology that is not locally convex.

2.2.6 Example: $L_p(\Omega, \Sigma, \mu)$, $0 < p < 1$. Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω and that $0 < p < 1$. Define $L_p(\Omega, \Sigma, \mu)$ to be the collection of all μ-measurable scalar-valued functions on Ω such that $\int_\Omega |f|^p \, d\mu$ is finite, with the usual convention that functions that agree almost everywhere are considered to be the same. If $s \ge 0$ and $\phi_s(t) = s^p + t^p - (s+t)^p$ whenever $t \ge 0$, then ϕ_s is nondecreasing and hence nonnegative on $[0, \infty)$, from which it follows that

$$|f(x) + g(x)|^p \le \big(|f(x)| + |g(x)|\big)^p \le |f(x)|^p + |g(x)|^p$$

whenever $f, g \in L_p(\Omega, \Sigma, \mu)$ and $x \in \Omega$. Easy arguments based on this show that $L_p(\Omega, \Sigma, \mu)$ is a vector space under the usual addition of functions and multiplication of functions by scalars, and that the formula

$$d(f,g) = \int_\Omega |f - g|^p \, d\mu$$

defines a metric on $L_p(\Omega, \Sigma, \mu)$. *Henceforth, whenever* $0 < p < 1$ *and* $L_p(\Omega, \Sigma, \mu)$ *is treated as a topological space, the topology is assumed to be the one induced by this metric.*

The next order of business is to show that the metric d is complete. If $f, g \in L_p(\Omega, \Sigma, \mu)$ and $\epsilon > 0$, then

$$\mu[\,|f-g| > \epsilon\,] = \int_{[\,|f-g|>\epsilon\,]} 1 \, d\mu \le \epsilon^{-p} \int_{[\,|f-g|>\epsilon\,]} |f-g|^p \, d\mu \le \epsilon^{-p} d(f,g),$$

from which it follows that every Cauchy sequence in $L_p(\Omega, \Sigma, \mu)$ is Cauchy in measure and so is convergent in measure to some measurable function. Let (f_n) be a Cauchy sequence in $L_p(\Omega, \Sigma, \mu)$ and let f be the function to which (f_n) converges in measure. It is a standard fact from measure theory that there is a subsequence (f_{n_j}) of (f_n) that converges to f almost everywhere. Suppose that $\epsilon > 0$. Let the positive integer N_ϵ be such that

$d(f_m, f_n) < \epsilon$ whenever $m, n \geq N_\epsilon$. If $n \geq N_\epsilon$, then an application of Fatou's lemma shows that

$$\int_\Omega |f - f_n|^p \, d\mu = \int_\Omega \lim_j |f_{n_j}(x) - f_n(x)|^p \, d\mu(x) \leq \liminf_j d(f_{n_j}, f_n) \leq \epsilon.$$

It follows that $f = (f - f_{N_\epsilon}) + f_{N_\epsilon} \in L_p(\Omega, \Sigma, \mu)$ and that $d(f, f_n) \to 0$, which establishes the completeness of d.

The metric d is invariant, so vector addition is continuous for $L_p(\Omega, \Sigma, \mu)$ since the space is a topological group under this operation. Furthermore, if (α_n) and (f_n) are sequences in $\mathbb{F}$ and $L_p(\Omega, \Sigma, \mu)$ respectively and have the respective limits α and f, then

$$\begin{aligned} d(\alpha_n f_n, \alpha f) &\leq d(\alpha_n f_n, \alpha_n f) + d(\alpha_n f, \alpha f) \\ &= |\alpha_n|^p \int_\Omega |f_n - f|^p \, d\mu + |\alpha_n - \alpha|^p \int_\Omega |f|^p \, d\mu \to 0 \end{aligned}$$

as $n \to \infty$, so multiplication of vectors by scalars is also a continuous operation. Thus, the metric d induces a vector topology on $L_p(\Omega, \Sigma, \mu)$.

Specializing to the case of Lebesgue measure λ on the Lebesgue-measurable subsets of $[0,1]$ yields the space $L_p[0,1]$. Notice that $L_p[0,1]$ contains every member of $L_\infty[0,1]$ and so has nonzero elements. Suppose that C is a convex neighborhood of 0 in $L_p[0,1]$ and that ϵ is a positive number such that the open ball of radius ϵ centered at 0 is included in C. Let f be a member of $L_p[0,1]$ and let n be a positive integer large enough that $d(f, 0) < n^{1-p}\epsilon$. Since the function $t \mapsto \int_{[0,t]} |f|^p \, d\lambda$ is continuous on $[0,1]$, there must be numbers $t_0, \ldots, t_n$ such that $0 = t_0 < t_1 < \cdots < t_n = 1$ and

$$\int_{[t_{j-1}, t_j)} |f|^p \, d\lambda = \frac{1}{n} \int_{[0,1]} |f|^p \, d\lambda$$

for each integer j such that $1 \leq j \leq n$. For each such j, let $\mathbf{I}_{[t_{j-1}, t_j)}$ be the indicator function of the interval $[t_{j-1}, t_j)$ and let the member f_j of $L_p[0,1]$ be defined by the formula $f_j(t) = nf(t)\mathbf{I}_{[t_{j-1}, t_j)}(t)$. Then

$$d(f_j, 0) = n^p \int_{[t_{j-1}, t_j)} |f|^p \, d\lambda = n^{p-1} \int_{[0,1]} |f|^p \, d\lambda = n^{p-1} d(f, 0) < \epsilon$$

for each j, so each f_j is in C. Since $f = \sum_{j=1}^n n^{-1} f_j$, the convex combination f of members of C is itself in C, so $C = L_p[0,1]$. Since every nonempty convex open subset of $L_p[0,1]$ is, by Proposition 2.1.39 (b), a translate of a convex neighborhood of 0, the only nonempty convex open subset of $L_p[0,1]$ is $L_p[0,1]$ itself. Thus, the topological vector space $L_p[0,1]$ is not locally convex.

2.2.7 Example: ℓ_p, $0 < p < 1$. Let p be a real number such that $0 < p < 1$ and let μ be the counting measure on the collection Σ of all subsets of $\mathbb{N}$.

By analogy with what is done when $1 \le p \le \infty$, the space ℓ_p is defined to be $L_p(\mathbb{N}, \Sigma, \mu)$. Notice that ℓ_p is the space of all sequences (α_n) such that $\sum_n |\alpha_n|^p$ is finite, and that if $(\alpha_n), (\beta_n) \in \ell_p$, then

$$d\big((\alpha_n), (\beta_n)\big) = \sum_n |\alpha_n - \beta_n|^p.$$

Suppose that C is a convex neighborhood of 0 in ℓ_p and that ϵ is such that the open ball of radius 2ϵ centered at 0 is included in C. Let $\{\, e_n : n \in \mathbb{N} \,\}$ be the collection of standard coordinate vectors of ℓ_p. Then $\epsilon^{1/p} e_n \in C$ for each n, so $m^{-1}\epsilon^{1/p} \sum_{j=1}^m e_j \in C$ for each positive integer m. However,

$$d\Big(\frac{\epsilon^{1/p}}{m} \sum_{j=1}^m e_j, 0\Big) = \sum_{j=1}^m \left|\frac{\epsilon^{1/p}}{m}\right|^p = \epsilon m^{1-p} \to \infty$$

as $m \to \infty$, so C is unbounded with respect to the metric d. Thus, no open ball centered at 0 includes a convex neighborhood of 0, and so the metric of ℓ_p is another example of a complete invariant metric inducing a vector topology that is not locally convex.

The notion of boundedness with respect to a metric, which played such an important role in the preceding example, is so useful that it would be good to be able to extend it somehow to vector topologies not induced by metrics. One reasonable approach would be to call a set A bounded if the points of A are "uniformly close enough to 0" that $tx \to 0$ "uniformly" for x in A as t decreases to 0; that is, if the following happens: For each neighborhood U of 0, there is a positive r_U such that $tA \subseteq U$ whenever $0 < t < r_U$. This requirement is clearly equivalent to the condition in the following definition.

2.2.8 Definition. A subset A of a TVS is *bounded* if, for each neighborhood U of 0, there is a positive s_U such that $A \subseteq tU$ whenever $t > s_U$.

A moment's thought shows that a subset of a normed space is bounded in the sense of the preceding definition if and only if it is bounded with respect to the metric induced by the norm. However, it can happen that a vector topology is induced by a metric d such that the metrically bounded sets are not the same as the sets bounded in the sense of Definition 2.2.8. As an example, consider the metric d on $\mathbb{R}$ defined by the formula

$$d(x, y) = \min\{1, |x - y|\},$$

a complete invariant metric that induces the usual norm topology of $\mathbb{R}$. Notice that $\mathbb{R}$ is itself metrically bounded with respect to d, though not bounded in the sense of Definition 2.2.8. *Henceforth, the term "bounded"* *when applied to a TVS is always to be interpreted in the sense of Definition 2.2.8 unless specifically stated otherwise.*

One reason that a TVS is a good place to do analysis is that many familiar properties of normed spaces, and in particular of $\mathbb{R}$ and $\mathbb{C}$, generalize to TVSs. Here is a list of some of the most useful ones.

2.2.9 Theorem. *Suppose that X is a topological vector space.*

(a) *Let (β_α) be a net in $\mathbb{F}$ and let (x_α) and (y_α) be nets in X such that all three nets have the same index set and (β_α), (x_α), and (y_α) converge to β, x, and y respectively. Let γ and z be elements of $\mathbb{F}$ and X respectively. Then $x_\alpha + y_\alpha \to x + y$, $\beta_\alpha x_\alpha \to \beta x$, $x_\alpha + z \to x + z$, $\gamma x_\alpha \to \gamma x$, and $\beta_\alpha z \to \beta z$.*

(b) *If f and g are continuous functions from a topological space into X and α is a scalar, then $f + g$ and αf are continuous.*

(c) *Let x_0 be an element of X and let α_0 be a nonzero scalar. Then the maps $x \mapsto x + x_0$ and $x \mapsto \alpha_0 x$ are homeomorphisms from X onto itself. Consequently, if A is a subset of X that is open, or closed, or compact, then $x_0 + A$ and $\alpha_0 A$ also have that property. If A and U are subsets of X and U is open, then $A + U$ is open.*

(d) *Suppose that A and B are subsets of X, that $x_0 \in X$, and that α_0 is a nonzero scalar. Then $\overline{A}+\overline{B} \subseteq \overline{A + B}$, $x_0+\overline{A} = \overline{x_0 + A}$, $\alpha_0\overline{A} = \overline{\alpha_0 A}$, $A^\circ + B^\circ \subseteq (A + B)^\circ$, $x_0 + A^\circ = (x_0 + A)^\circ$, and $\alpha_0 A^\circ = (\alpha_0 A)^\circ$.*

(e) *For each x_0 in X, the neighborhoods of x_0 are exactly the sets x_0+U such that U is a neighborhood of 0.*

(f) *Each neighborhood of 0 in X is absorbing.*

(g) *For each neighborhood U of 0 in X, there is a balanced neighborhood V of 0 in X such that $V \subseteq \overline{V} \subseteq \overline{V+V} \subseteq U$. If U is convex, then V can be chosen to be convex.*

(h) *Suppose that A is a bounded subset of X, that x_0 is an element of X, and that α_0 is a scalar. Then $x_0 + A$ and $\alpha_0 A$ are bounded.*

(i) *Let A be a subset of X. Then $[A] = \overline{\langle A\rangle}$ and $\overline{\text{co}}(A) = \overline{\text{co}(A)}$. If A is a subspace of X, then so is $\overline{A}$. If A is balanced, then so is $\overline{A}$, and A° is also balanced provided that $0 \in A^\circ$. If A is bounded or convex, then both $\overline{A}$ and A° have that same property.*

(j) *Let Y be a subspace of X. Then the relative topology that Y inherits from X is a vector topology. If the topology of X is locally convex, then so is the relative topology of Y.*

PROOF. Part (a) is just a collection of special cases of Corollary 2.1.23. Part (b) is an easy consequence of the continuity of the vector space operations of X, and in particular follows readily from (a). For (c), suppose that $x_0 \in X$ and α_0 is a nonzero scalar. The continuity of the vector space operations of X implies that the maps $x \mapsto x+x_0$ and $x \mapsto \alpha_0 x$ are continuous, as are their respective inverses $x \mapsto x - x_0$ and $x \mapsto \alpha_0^{-1}x$, so the two maps of part (c) are homeomorphisms from X onto itself. If A is a subset of X that is open, or closed, or compact, then $x_0 + A$ and $\alpha_0 A$ have that same property since these properties are preserved by homeomorphisms. If A and U are subsets of X and U is open, then $A+U$ is open as the union of open sets, since $A + U = \bigcup\{ a + U : a \in A \}$. This proves (c).

Suppose that A, B, x_0, and α_0 are as in the hypotheses of (d). Let y_0 and z_0 be elements of $\overline{A}$ and $\overline{B}$ respectively, and let (y_α) and (z_β) be nets in A and B respectively such that $y_\alpha \to y_0$ and $z_\beta \to z_0$. There are subnets (y_γ) and (z_γ) of (y_α) and (z_β) respectively having the same index set; see Technique 2.1.32. Then $y_\gamma + z_\gamma \to y_0 + z_0$, so $y_0 + z_0 \in \overline{A+B}$. This proves that $\overline{A} + \overline{B} \subseteq \overline{A+B}$. Since $A^\circ + B^\circ$ is an open subset of $A + B$, it follows that $A^\circ + B^\circ \subseteq (A+B)^\circ$. Straightforward arguments based on (a) and (c) yield the rest of (d).

Part (e) follows immediately from Proposition 2.1.39 (b) and the fact that X is a topological group under addition of vectors. For (f), suppose that U is a neighborhood of 0 in X. If $x \in X$, then the continuity of multiplication of vectors by scalars yields a positive r_x such that $tx \in U$ when $0 < t < r_x$, so $x \in tU$ when $t > r_x^{-1}$. Thus, the set U is absorbing, which proves (f).

To prove (g), suppose that U is a neighborhood of 0 in X. The continuity of vector addition yields neighborhoods U_1 and U_2 of 0 such that $U_1 + U_2 \subseteq U$. Let $U_3 = U_1 \cap U_2 \cap (-U_1) \cap (-U_2)$. Then U_3 is a neighborhood of 0 such that $U_3 = -U_3$ and $U_3 + U_3 \subseteq U$. The same procedure applied to U_3 instead of U yields a neighborhood U_4 of 0 such that $U_4 = -U_4$ and $U_4 + U_4 + U_4 + U_4 \subseteq U$. It follows that $U_4 + U_4 + U_4$ does not intersect $X \setminus U$, so the fact that $U_4 = -U_4$ implies that $U_4 + U_4$ does not intersect $(X \setminus U) + U_4$. Since $(X \setminus U) + U_4$ is open, it follows that $\overline{U_4 + U_4}$ does not intersect $(X \setminus U) + U_4$, so $\overline{U_4 + U_4} \subseteq U$. The continuity of multiplication of vectors by scalars produces a positive δ and a neighborhood U_5 of 0 in X such that $\alpha U_5 \subseteq U_4$ whenever $|\alpha| < \delta$. Let $V = \bigcup\{ \alpha U_5 : |\alpha| < \delta \}$. Then V is a balanced neighborhood of 0 lying in U_4, and

$$V \subseteq \overline{V} = \overline{V + \{0\}} \subseteq \overline{V + V} \subseteq \overline{U_4 + U_4} \subseteq U.$$

This proves all of (g) except for the convexity assertion, which uses (i) in its proof and will be obtained below.

Suppose that A is a bounded subset of X, that x_0 is an element of X, and that α_0 is a scalar. Let U be a neighborhood of 0 in X and let V be a balanced neighborhood of 0 in X such that $V+V \subseteq U$. Let s_V be a positive number such that $A \subseteq tV$ and $x_0 \in tV$ whenever $t > s_V$. If $t > s_V$, then $x_0 + A \subseteq tV + tV \subseteq tU$, from which it follows that $x_0 + A$ is bounded. If $t > \big(|\alpha_0| + 1\big)s_V$, then $\alpha_0 A \subseteq t\big(|\alpha_0| + 1\big)^{-1}\alpha_0 V \subseteq tV \subseteq tU$, which shows that $\alpha_0 A$ is bounded and finishes the proof of (h).

Let A be a subset of X. Suppose for the remainder of this paragraph that $x, y \in \overline{A}$ and $\alpha, \beta \in \mathbb{F}$. As in the proof of (d), it is possible to find nets (x_γ) and (y_γ) in A having the same index set such that $x_\gamma \to x$ and $y_\gamma \to y$. Then $\alpha x_\gamma + \beta y_\gamma \to \alpha x + \beta y$. If $\alpha x_\gamma + \beta y_\gamma \in A$ for each γ, then $\alpha x + \beta y \in \overline{A}$. It follows that $\overline{A}$ is a subspace of X if A is a subspace of X and, by considering the case in which $\alpha, \beta > 0$ and $\alpha + \beta = 1$, that $\overline{A}$ is convex if A is convex. The claim in (i) about linear hulls and convex hulls now follows easily; see the proof of Corollary 1.3.11. If A is balanced and $|\alpha| \leq 1$, then each αx_γ is in A, so $\alpha x \in \overline{A}$ since $\alpha x_\gamma \to \alpha x$, which implies that $\overline{A}$ is balanced. Now suppose that A is bounded. Let U be a neighborhood of 0 in X and let V be a neighborhood of 0 in X such that $\overline{V} \subseteq U$. Let s_V be a positive number such that $A \subseteq tV$ when $t > s_V$. Then $\overline{A} \subseteq \overline{tV} = t\overline{V} \subseteq tU$ when $t > s_V$, so $\overline{A}$ is bounded. This proves all the assertions about closures in (i).

If A is a convex subset of X and $0 < t < 1$, then $tA^\circ + (1-t)A^\circ \subseteq A$, which implies that the open set $tA^\circ + (1-t)A^\circ$ is included in A° and therefore that A° is convex. If A is a bounded subset of X, then its subset A° is clearly bounded. Now suppose that A is a balanced subset of X such that $0 \in A^\circ$. Let α be a scalar such that $0 < |\alpha| \leq 1$. Then $\alpha A^\circ \subseteq \alpha A \subseteq A$, so $\alpha A^\circ \subseteq A^\circ$ since αA° is open. It is also true that $0A^\circ = \{0\} \subseteq A^\circ$, so A° is balanced. This finishes the proof of (i).

Let Y be a subspace of X. The vector space operations of Y inherit the continuity of those of X, so the relative topology of Y is a vector topology. If the topology of X has a basis $\mathfrak{B}$ consisting of convex sets, then $\{ B \cap Y : B \in \mathfrak{B} \}$ is a basis for the relative topology of Y consisting of convex sets, so the relative topology of Y is locally convex. This proves (j).

All that remains to be proved is the convexity assertion in (g). Let U be a convex neighborhood of 0 in X. It suffices to find a set V with the appropriate properties inside the convex neighborhood $U \cap (-U)$ of 0, so it can be assumed that $U = -U$. Since $3^{-1}U + 3^{-1}U - 3^{-1}U = U$, it follows that $3^{-1}U + 3^{-1}U$ does not intersect the open set $(X \setminus U) + 3^{-1}U$ that includes $X \setminus U$, so $\overline{3^{-1}U + 3^{-1}U} \subseteq U$. It is enough to find a convex balanced neighborhood V of 0 such that $V \subseteq 3^{-1}U$. Let

$$W = \bigcap \{ 3^{-1}\alpha U : \alpha \in \mathbb{F},\ |\alpha| = 1 \}.$$

Then W is a subset of $3^{-1}U$ and is convex as the intersection of convex sets. Let B be a balanced neighborhood of 0 included in $3^{-1}U$. If α is a scalar such that $|\alpha| = 1$, then $B = \alpha B \subseteq 3^{-1}\alpha U$, and so $B \subseteq W$, from which it follows that $0 \in W^\circ$. Now let β be a scalar such that $|\beta| \leq 1$ and let t and γ be scalars such that $0 \leq t \leq 1$, $|\gamma| = 1$, and $\beta = t\gamma$. Then

$$\beta W = t\Big(\bigcap \{3^{-1}\alpha\gamma U : \alpha \in \mathbb{F},\ |\alpha| = 1 \}\Big) = tW = tW + (1-t)\{0\} \subseteq W,$$

so W is balanced. It follows from (i) that W° is a convex balanced neighborhood of 0 included in $3^{-1}U$, and so is the desired V. ■

2.2.10 Corollary. *A subset A of a TVS X is bounded if and only if it has this property: For each balanced neighborhood U of 0 in X, there is a positive s_U such that $A \subseteq s_U U$.*

PROOF. The "only if" portion follows immediately from the definition of boundedness. For the "if" portion, suppose that A has the indicated property for each balanced neighborhood of 0 in X. Let U be a neighborhood of 0 in X and let V be a balanced neighborhood of 0 included in U. Let s_V be a positive number such that $A \subseteq s_V V$. If $t > s_V$, then

$$A \subseteq s_V V = t(t^{-1}s_V V) \subseteq tV \subseteq tU.$$

It follows that A is bounded. ■

If U is a neighborhood of 0 in $\mathbb{R}$, then U includes some balanced neighborhood $(-\epsilon, \epsilon)$ of 0, and so there is a balanced neighborhood $(-\epsilon/2, \epsilon/2)$ of 0 such that $(-\epsilon/2, \epsilon/2) + (-\epsilon/2, \epsilon/2) \subseteq U$. Therefore, if $x, y \in \mathbb{R}$ and $|x|, |y| < \epsilon/2$, which is to say that $x, y \in (-\epsilon/2, \epsilon/2)$, then $x + y \in U$. Many familiar "epsilon over two" arguments of analysis are based on this. Such arguments can often be generalized to TVSs because of part (g) of the preceding theorem. See Exercise 2.27 for an example.

It is often easier to use Corollary 2.2.10 to check for boundedness than to proceed directly from the definition. The proof of the next result gives an example of this.

2.2.11 Proposition. *Every compact subset of a TVS is bounded. Thus, every convergent sequence in a TVS is bounded.*

PROOF. Let K be a compact subset of a TVS X and let U be a balanced neighborhood of 0 in X. Since U is absorbing, the collection $\{ tU : t > 0 \}$ is an open covering for K, so there are positive numbers $t_1, \ldots, t_n$ such that $t_1 < t_2 < \cdots < t_n$ and $K \subseteq \bigcup_{j=1}^{n} t_j U$. Since $t_j U = t_n(t_n^{-1} t_j U) \subseteq t_n U$ for each j, it follows that $K \subseteq t_n U$, so Corollary 2.2.10 implies that K is bounded. The rest of the proposition then follows from the fact that a set in a topological space consisting of the terms and a limit of a convergent sequence must be compact. ■

One fairly obvious convention is needed before proceeding. *Whenever terminology from the theory of abelian topological groups is used in a situation involving topological vector spaces, the TVSs in question are being viewed as abelian topological groups under vector addition.* In particular, a net $(x_\alpha)_{\alpha \in I}$ in a TVS X is *Cauchy* if, for every neighborhood U of 0 in X, there is an α_U in I such that $x_\beta - x_\gamma \in U$ whenever $\alpha_U \preceq \beta$ and $\alpha_U \preceq \gamma$, and X is *complete* if each Cauchy net in X converges; see Definitions 2.1.41 and 2.1.48.

By Proposition 2.1.47, every convergent net in a TVS is Cauchy. The fact that convergent sequences in a TVS are always bounded is therefore just a special case of the following more general result.

2.2.12 Proposition. *Every Cauchy sequence in a TVS is bounded.*

PROOF. Suppose that (x_n) is a Cauchy sequence in a TVS X. Let U be a neighborhood of 0 in X and let V be a balanced neighborhood of 0 such that $V + V \subseteq U$. Let N be a positive integer such that $x_n - x_N \in V$ whenever $n \geq N$. Since V is absorbing, there is a positive number s_N such that $x_N \in s_N V$. It follows that if $n \geq N$ and $t > \max\{1, s_N\}$, then

$$x_n = (x_n - x_N) + x_N \in V + s_N V \subseteq tV + tV \subseteq tU.$$

Finally, let $\{s_1, \ldots, s_{N-1}\}$ be positive numbers such that $x_j \in tU$ when $1 \leq j \leq N-1$ and $t > s_j$, and let $s = \max\{1, s_1, \ldots, s_{N-1}, s_N\}$. Then $\{\, x_n : n \in \mathbb{N} \,\} \subseteq tU$ whenever $t > s$, so the sequence (x_n) is bounded. ■

The separation properties of the open subsets of a TVS tend to be either very good or very bad. In fact, if X is a TVS that is not completely regular, then there is an entire nontrivial subspace Y of X such that every open set intersecting Y actually includes Y; see Exercise 2.33. This result follows from the fact, which will now be proved, that every T_0 vector topology is actually completely regular. The proof is based on the following construction, which will have a further application in the proof of Theorem 2.3.11.

2.2.13 Construction. Let B be a balanced neighborhood of 0 in a TVS X. A continuous nonnegative-real-valued function f on X will be constructed such that f is small "near the origin" and large "far from the origin," where the "distance" of a point from the origin is determined by how "far" the point is inside or outside B. Here are the steps in the construction.

1. Let $B(1) = B$. For each positive integer n, define $B(2^n)$ inductively by letting $B(2^n) = B(2^{n-1}) + B(2^{n-1})$. For each positive integer n, define $B(2^{-n})$ inductively by letting $B(2^{-n})$ be a balanced neighborhood of 0 such that $B(2^{-n}) + B(2^{-n}) \subseteq B(2^{-n+1})$.

2. The terminating binary expansion of the number $5\frac{9}{16}$ is 101.1001. Let

   $$B(5\tfrac{9}{16}) = 1B(4) + 0B(2) + 1B(1) + 1B(\tfrac{1}{2}) + 0B(\tfrac{1}{4}) + 0B(\tfrac{1}{8}) + 1B(\tfrac{1}{16}).$$

   Define $B(\frac{m}{2^n})$ analogously for each integer n and each positive integer m, that is, for each positive dyadic rational. Notice that if n is an integer, then the definition of $B(2^n)$ given in this step is consistent with that given in Step 1, and that the inclusion of extraneous leading or trailing zeros in the terminating binary expansion of a positive dyadic rational r does not affect the definition of $B(r)$.

3. **Claim:** $B(r_1) + B(r_2) \subseteq B(r_1 + r_2)$ whenever r_1 and r_2 are positive dyadic rationals. To see this, let r_1 and r_2 be positive dyadic rationals and let

   $$\delta^{(1)}_{n_2}\delta^{(1)}_{n_2-1}\cdots\delta^{(1)}_0.\delta^{(1)}_{-1}\cdots\delta^{(1)}_{n_1},$$
   $$\delta^{(2)}_{n_2}\delta^{(2)}_{n_2-1}\cdots\delta^{(2)}_0.\delta^{(2)}_{-1}\cdots\delta^{(2)}_{n_1},$$

and

$$\delta^{(3)}_{n_2}\delta^{(3)}_{n_2-1}\cdots\delta^{(3)}_{0}.\delta^{(3)}_{-1}\cdots\delta^{(3)}_{n_1}$$

be terminating binary expansions for r_1, r_2, and r_1+r_2 respectively, where extra leading and trailing zeros have been included where necessary to assure that all three expansions have the same number of digits to the left of the binary point and similarly for the number of digits to the right. The claim will be proved once it is shown that

$$\sum_{j=n_1}^{n_2}\delta^{(1)}_j B(2^j)+\sum_{j=n_1}^{n_2}\delta^{(2)}_j B(2^j)\subseteq\sum_{j=n_1}^{n_2}\delta^{(3)}_j B(2^j).$$

Let $A_{n_1}=\sum_{j=n_1}^{n_2}\delta^{(1)}_j B(2^j)+\sum_{j=n_1}^{n_2}\delta^{(2)}_j B(2^j)$ and use the following inductive process to pass from A_{n_1} to another set A_{n_2+1}. The sum representing A_{n_1} must have exactly zero, one, or two terms of the form $1B(2^{n_1})$. If it has two such terms, delete those two terms from the sum and insert a term of the form $1B(2^{n_1+1})$ to obtain A_{n_1+1}; otherwise, let $A_{n_1+1}=A_{n_1}$. It follows from Step 1 that $A_{n_1}\subseteq A_{n_1+1}$. The sum representing A_{n_1+1} must have exactly zero, one, two, *or three* terms of the form $1B(2^{n_1+1})$. If it has more than one such term, delete two of those terms from the sum and insert a term of the form $1B(2^{n_1+2})$ to obtain A_{n_1+2}; otherwise, let $A_{n_1+2}=A_{n_1+1}$. Then $A_{n_1+1}\subseteq A_{n_1+2}$. The sum representing A_{n_1+2} must have exactly zero, one, two, or three terms of the form $1B(2^{n_1+2})$. If it has more than one such term, delete two of those terms from the sum and insert a term of the form $1B(2^{n_1+3})$ to obtain A_{n_1+3}; otherwise, let $A_{n_1+3}=A_{n_1+2}$. Then $A_{n_1+2}\subseteq A_{n_1+3}$. Continue in the obvious fashion until A_{n_2+1} is obtained. Now think of how this inductive process is analogous to the usual algorithm for adding r_1 to r_2 by aligning their binary expansions one above the other by binary point and adding digits in columns, working from right to left and carrying a 1 to the left whenever necessary. It should be clear after a moment's thought that $A_{n_2+1}=\sum_{j=n_1}^{n_2}\delta^{(3)}_j B(2^j)$. Since $A_{n_1}\subseteq A_{n_2+1}$, the claim is proved.

4. **Claim:** If r_1 and r_2 are positive dyadic rationals such that $r_1<r_2$, then $B(r_1)\subseteq B(r_2)$. This follows from the preceding claim, since $0\in B(r_2-r_1)$ and $B(r_1)+B(r_2-r_1)\subseteq B(r_2)$.

5. For each positive real number t, let

$$B(t)=\bigcup\{\,B(r): r \text{ is a positive dyadic rational, } r\le t\,\}.$$

It follows from Step 4 that if t is a positive dyadic rational, then this definition of $B(t)$ is consistent with the definition from Step 2.

6. **Claims:** If $s, t > 0$, then $B(s) + B(t) \subseteq B(s+t)$, and if $0 < s < t$, then $B(s) \subseteq B(t)$. These claims follow in an obvious way from Steps 3 and 5.

7. **Claim:** For each positive t, the set $B(t)$ is a balanced neighborhood of 0. To see this, first observe that if U and V are balanced neighborhoods of 0, then so is $U + V$. It then follows from Step 1 that $B(2^n)$ is a balanced neighborhood of 0 for each integer n, from Step 2 that $B(r)$ is a balanced neighborhood of 0 for each positive dyadic rational r, and from Step 5 that $B(t)$ is a balanced neighborhood of 0 for each positive real number t.

8. **Claim:** $X = \bigcup\{\, B(t) : t > 0 \,\}$. To see this, suppose that $x \in X$. Let n be a positive integer such that $x \in 2^n B$. An easy argument based on Step 1 shows that $2^n B \subseteq \sum_{j=1}^{2^n} B = B(2^n)$, from which the claim follows.

9. For each x in X, let $f(x) = \inf\{\, t : t > 0,\, x \in B(t) \,\}$. It follows from Step 8 that f is finite-valued and so has its range in $[0, +\infty)$. It is clear that $f(0) = 0$ and that if $t > 0$, then $f(x) \le t$ if $x \in B(t)$ and $f(x) \ge t$ if $x \in X \setminus B(t)$. In particular, it follows that $f(x) \ge 1$ whenever $x \in X \setminus B$.

10. **Claim:** If $x \in X$, then $f(x) = f(-x)$. This is a straightforward consequence of the definition of f and the fact that $B(t) = -B(t)$ for each positive t.

11. **Claim:** If $x, y \in X$, then $f(x+y) \le f(x) + f(y)$ and $|f(x) - f(y)| \le f(x-y)$. To see this, suppose that $x, y \in X$ and that $\epsilon > 0$. Then there are positive reals s and t such that $s < f(x) + \epsilon$, $t < f(y) + \epsilon$, $x \in B(s)$, and $y \in B(t)$. Then $x + y \in B(s) + B(t) \subseteq B(s+t)$, and so $f(x+y) \le s + t < f(x) + f(y) + 2\epsilon$. It follows that $f(x+y) \le f(x) + f(y)$. Also,

$$\begin{aligned} f(x) - f(y) &= f(x - y + y) - f(y) \\ &\le f(x-y) + f(y) - f(y) \\ &= f(x-y), \end{aligned}$$

and similarly

$$f(y) - f(x) \le f(y-x) = f(x-y).$$

It follows that $|f(x) - f(y)| \le f(x-y)$, which finishes the proof of the claim.

12. **Claim:** The function f is continuous. For the proof of this, suppose that $x_0 \in X$ and that $\epsilon > 0$. If x is in the neighborhood $x_0 + B(\epsilon)$ of x_0, then $x - x_0 \in B(\epsilon)$, and so $|f(x) - f(x_0)| \le f(x - x_0) \le \epsilon$. It follows that f is continuous at x_0, which proves the claim and finishes the construction.

2.2.14 Theorem. *Every* T_0 *vector topology is completely regular.*

PROOF. Suppose that X is a TVS whose topology is T_0. The first order of business is to show that the topology is in fact T_1. Let x and y be distinct elements of X. Then there is a neighborhood U of 0 such that either $x \notin y+U$ or $y \notin x+U$; it can be assumed without loss of generality that $x \notin y+U$. Then y is not in the neighborhood $x-U$ of x, from which it follows that the topology of X is T_1.

Now let x_0 be an element of X and let F be a closed subset of X not containing x_0. Since 0 is not in the closed set $-x_0+F$, there is a balanced neighborhood B of 0 such that $B\cap(-x_0+F) = \varnothing$. By Construction 2.2.13, there is a continuous function $f\colon X \to [0,+\infty)$ such that $f(0) = 0$ and $f(x) \geq 1$ whenever $x \in X \setminus B$; notice that $f(x) \geq 1$ for each x in $-x_0+F$. Let $g(x) = \min\{1, f(x-x_0)\}$ whenever $x \in X$. Then g is a continuous function from X into $[0,1]$ such that $g(x_0) = 0$ and $g(x) = 1$ whenever $x \in F$. The topology of X is therefore completely regular. ■

Thus, a vector topology that satisfies any of the separation axioms T_0 through $T_{3\frac{1}{2}}$ actually satisfies all of them. It is traditional that such vector topologies be called Hausdorff, but it should be kept in mind that for vector topologies the Hausdorff axiom is implied by the T_0 axiom and implies complete regularity.

Suppose that X is a vector space with a topology of any sort. If x_1^* and x_2^* are continuous linear functionals on X and α is a scalar, then the continuity of $\alpha x_1^* + x_2^*$ follows easily from the continuity of addition and multiplication in $\mathbb{F}$; see also Corollary 1.3.4. The collection of all continuous linear functionals on X therefore forms a subspace of the vector space $X^\#$ of all linear functionals on X, so the following definition is justified.

2.2.15 Definition. Suppose that X is a vector space with a topology. The *dual space* of X or *dual* of X or *conjugate* of X, denoted by X^*, is the vector space of all continuous linear functionals on X with the obvious vector space operations.

There is a slight technical conflict between this definition and the one previously given for the dual space of a normed space. If X is a normed space, then by Definition 2.2.15 the dual space of X should be the *vector* space of continuous linear functionals on X, while by Definition 1.10.1 it should be the corresponding *normed* space. The only time this could cause any real problem is when the vector space X underlying a normed space is given a topology $\mathfrak{T}$ different from its usual norm topology, and "X^*" could refer to either $(X, \|\cdot\|)^*$ or $(X,\mathfrak{T})^*$, dual spaces that might not even have the same underlying vector space; see Exercise 2.30. In this situation, care will be taken to assure that no confusion results.

The following theorem, an extension to topological vector spaces of parts of Theorems 1.4.2 and 1.7.15 and Proposition 1.7.16, provides several ways to test a linear functional on a TVS for continuity.

2.2.16 Theorem. *Suppose that x^* is a linear functional on a TVS X. Then the following are equivalent.*

(a) *The functional x^* is continuous.*

(b) *There is a neighborhood U of 0 in X such that $x^*(U)$ is a bounded subset of $\mathbb{F}$.*

(c) *The kernel of x^* is a closed subset of X.*

(d) *The kernel of x^* is not a proper dense subset of X.*

PROOF. The theorem is trivially true if x^* is the zero functional, so it can be assumed that it is not. Since the kernel of x^* is the inverse image under x^* of the closed subset $\{0\}$ of $\mathbb{F}$, it is clear that (a) $\Rightarrow$ (c) $\Rightarrow$ (d). Suppose that (d) holds. Fix an element x_0 of $X \setminus \overline{\ker(x^*)}$. Then parts (e) and (g) of Theorem 2.2.9 together imply that there is a balanced neighborhood U of 0 such that $x_0 + U \subseteq X \setminus \overline{\ker(x^*)}$. Notice that $x^*u \neq -x^*x_0$ whenever $u \in U$. Since the balanced subset $x^*(U)$ of $\mathbb{F}$ contains with each of its members every scalar of smaller absolute value, it must omit every scalar having absolute value larger than $|-x^*x_0|$, and so must be bounded. This proves that (d) $\Rightarrow$ (b).

Finally, suppose that there is a neighborhood U of 0 in X such that $x^*(U)$ is bounded. By multiplying U by a positive scalar if necessary, it may be assumed that $|x^*x| < 1$ whenever $x \in U$, and therefore that $|x^*x| < \epsilon$ whenever $\epsilon > 0$ and $x \in \epsilon U$. If $x \in X$ and $\epsilon > 0$, then $|x^*y - x^*x| = |x^*(y - x)| < \epsilon$ whenever $y \in x + \epsilon U$, from which it follows that x^* is continuous. This shows that (b) $\Rightarrow$ (a) and finishes the proof. ■

It is often important to know when two convex subsets C_1 and C_2 of a TVS X can be separated, not by open sets or by a continuous function in the sense of the topological separation axioms, but rather by a member x^* of X^* in the sense that $\sup\{\operatorname{Re} x^*x : x \in C_1\} \leq \inf\{\operatorname{Re} x^*x : x \in C_2\}$. Results of this type include Proposition 1.9.15 as well as the three separation theorems about to be proved here, and are often called collectively the *Hahn-Banach separation theorem* since they tend to be straightforward corollaries of the vector space version of the Hahn-Banach extension theorem.

2.2.17 Definition. Let X be a vector space. A *flat* or *affine* subset of X is a translate of a subspace of X, that is, a set of the form $x + Y$ where $x \in X$ and Y is a subspace of X.

2.2.18 Lemma. *Suppose that C is a convex subset of a TVS X. If $x \in C$, $y \in C^\circ$, and $0 < t < 1$, then $tx + (1 - t)y \in C^\circ$.*

PROOF. Just notice that $tx + (1-t)y \in tC + (1-t)C^\circ \subseteq C$ and that $tC + (1-t)C^\circ$ is an open set. ■

2.2.19 Mazur's Separation Theorem. (G. Ascoli, 1932 [8]; S. Mazur, 1933 [162]; D. G. Bourgin, 1943 [36]). *Let X be a TVS and let F and C be subsets of X such that F is flat and C is convex with nonempty interior. If $F \cap C^\circ = \varnothing$, then there is an x^* in X^* and a real number s such that*

(1) $\operatorname{Re} x^*x = s$ *for each x in F;*

(2) $\operatorname{Re} x^*x \le s$ *for each x in C; and*

(3) $\operatorname{Re} x^*x < s$ *for each x in C°.*

PROOF. Suppose first that the scalar field is $\mathbb{R}$ and that $0 \in C^\circ$. Then C is a convex absorbing subset of X, so by Proposition 1.9.14 the Minkowski functional p of C is sublinear and

$$\{\, x : x \in X,\, p(x) < 1 \,\} \subseteq C \subseteq \{\, x : x \in X,\, p(x) \le 1 \,\}.$$

The continuity of multiplication of vectors by scalars implies that for each x in C° there is an s_x such that $s_x > 1$ and $s_x x \in C^\circ$, so that $s_x p(x) = p(s_x x) \le 1$. It follows that $p(x) < 1$ whenever $x \in C^\circ$. Conversely, suppose that $x \in X$ and $p(x) < 1$. Then there is a t_x such that $t_x > 1$ and $p(t_x x) = t_x p(x) < 1$, which implies that $t_x x \in C$. Since $x = t_x^{-1}(t_x x) + (1 - t_x^{-1})0$, the lemma implies that $x \in C^\circ$. It follows that $C^\circ = \{\, x : x \in X,\, p(x) < 1 \,\}$.

Let Y be a subspace of X and x_0 an element of X such that $F = x_0 + Y$. Since $0 \notin F$, the subspace Y contains neither $-x_0$ nor its negative x_0, from which it follows that each element of the subspace $Y + \langle\{x_0\}\rangle$ of X has a unique representation of the form $y + \alpha x_0$ where $y \in Y$ and $\alpha \in \mathbb{R}$. Let $x_0^*(y + \alpha x_0) = \alpha$ whenever $y \in Y$ and $\alpha \in \mathbb{R}$. Then x_0^* is a linear functional on $Y + \langle\{x_0\}\rangle$. If α is a positive scalar and $y \in Y$, then $\alpha^{-1}y + x_0$ is in F and so is not in C°, from which it follows that

$$x_0^*(y + \alpha x_0) = \alpha \le \alpha p(\alpha^{-1}y + x_0) = p(y + \alpha x_0).$$

Since $p(x) \ge 0$ for each x in X, it is also true that $x_0^*(y + \alpha x_0) \le p(y + \alpha x_0)$ whenever $y \in Y$ and $\alpha \le 0$, so x_0^* is dominated by p on $Y + \langle\{x_0\}\rangle$. By the vector space version of the Hahn-Banach extension theorem, the functional x_0^* can be extended to a linear functional x^* on X such that $x^*x \le p(x)$ whenever $x \in X$. Now C° includes a balanced neighborhood U of 0, and $x^*(U)$ is a bounded subset of $\mathbb{R}$ since

$$|x^*u| = \max\{x^*(-u), x^*u\} \le \max\{p(-u), p(u)\} < 1$$

whenever $u \in U$. An application of Theorem 2.2.16 shows that $x^* \in X^*$. Since x^* is dominated by p, it follows that $x^*x \le 1$ when $x \in C$ and that $x^*x < 1$ when $x \in C^\circ$. Since $F = x_0 + Y$, it follows that $x^*x = x_0^*x = 1$ when $x \in F$, so this x^* satisfies the conclusion of the theorem when $s = 1$.

Now drop the assumption that $0 \in C^\circ$. Let x_1 be an element of C°. Then the interior $-x_1 + C^\circ$ of the convex set $-x_1 + C$ contains 0 and does not intersect the flat subset $-x_1 + F$ of X, so there is an x^* in X^* such that $x^*x = 1$ when $x \in -x_1+F$, $x^*x \le 1$ when $x \in -x_1+C$, and $x^*x < 1$ when $x \in -x_1+C^\circ$. It follows that $x^*x = x^*x_1 + 1$ when $x \in F$, $x^*x \le x^*x_1+1$ when $x \in C$, and $x^*x < x^*x_1 + 1$ when $x \in C^\circ$, and again the conclusion of the theorem holds.

Finally, suppose that the scalar field is $\mathbb{C}$. Let X_r be the real TVS obtained by restricting multiplication of vectors by scalars to $\mathbb{R} \times X$. Since every subspace of X is also a subspace of X_r, the set F is flat in X_r. It follows that there is a continuous real-linear functional z^* on X and a real number s such that $z^*x = s$ when $x \in F$, $z^*x \le s$ when $x \in C$, and $z^*x < s$ when $x \in C^\circ$. Let $x^*x = z^*x - iz^*(ix)$ for each x in X. It follows from Proposition 1.9.3 that x^* is a complex-linear functional on X with real part z^*. The continuity of z^* and of the vector space operations of X and $\mathbb{C}$ implies that $x^* \in X^*$, so x^* has all the required properties. ■

The following three corollaries are analogs for locally convex spaces of the normed space version of the Hahn-Banach extension theorem and of Corollaries 1.9.7 and 1.9.9 of that theorem. Notice that the third one requires the topology to be Hausdorff.

2.2.20 Corollary. *Let Y be a closed subspace of an LCS X. Suppose that $x \in X \setminus Y$. Then there is an x^* in X^* such that $x^*x = 1$ and $Y \subseteq \ker(x^*)$.*

PROOF. Since $X \setminus Y$ is a neighborhood of x, there is a convex neighborhood C of x that does not intersect the flat subset Y of X. The theorem yields an x_0^* in X^* and a real number s such that $\operatorname{Re} x_0^*z < s$ when $z \in C$ and $\operatorname{Re} x_0^*y = s$ when $y \in Y$. Since $0 \in Y$, it follows that $s = 0$, and therefore that x_0^*y (which equals $\operatorname{Re} x_0^*y - i \operatorname{Re} x_0^*(iy)$ if $\mathbb{F} = \mathbb{C}$) is 0 when $y \in Y$. Let $x^* = (x_0^*x)^{-1}x_0^*$. Then $x^*x = 1$ and $\ker(x^*) = \ker(x_0^*) \supseteq Y$, which finishes the proof. ■

2.2.21 Corollary. *Suppose that Y is a subspace of an LCS X and that $y^* \in Y^*$. Then there is an x^* in X^* whose restriction to Y is y^*.*

PROOF. The zero element of X^* extends the zero element of Y^* to X, so it can be assumed that $y^* \neq 0$ and therefore that there is a y_0 in Y such that $y^*y_0 = 1$. Let $Z = \overline{\ker(y^*)}$, where the closure is taken in X. The continuity of y^* and the fact that the topology of Y is inherited from X together imply that $y_0 \notin Z$, so by the preceding corollary there is an x^* in X^* such that $x^*y_0 = 1$ and $Z \subseteq \ker(x^*)$. If $y \in Y$, then $(y^*y)y_0 - y$ is in $\ker(y^*)$ and so in $\ker(x^*)$, which implies that

$$x^*y = x^*y + x^*\big((y^*y)y_0 - y\big) = x^*y + (y^*y)(x^*y_0) - x^*y = y^*y.$$

The restriction of x^* to Y is therefore y^*. ■

2.2.22 Corollary. *If x and y are different elements of a Hausdorff LCS X, then there is an x^* in X^* such that $x^*x \neq x^*y$.*

PROOF. Since $x - y$ is not in the closed subspace $\{0\}$ of X, Corollary 2.2.20 produces an x^* in X^* such that $x^*x - x^*y = x^*(x - y) = 1$. ■

The conclusion of the preceding corollary does not in general hold if the vector topology is not required to be Hausdorff or is not required to be locally convex. The examples about to be given to illustrate this depend on the fact that if X is a TVS whose only convex open subsets are the empty set and X itself, then X^* contains only the zero functional. To see this, suppose that X is a TVS having only those two convex open subsets and that $x^* \in X^*$. For each open ball U centered at the origin of $\mathbb{F}$, the subset $(x^*)^{-1}(U)$ of X is a nonempty convex open subset of X and so must be X. Since $x^*(X)$ is either $\{0\}$ or $\mathbb{F}$, it follows that $x^* = 0$.

2.2.23 Example. This is a continuation of Example 2.2.4. Let X be a vector space having a nonzero vector. Then the topology $\{\varnothing, X\}$ makes X into an LCS that is not Hausdorff. Since X^* contains only the zero functional, it is not possible to separate two different vectors of X by a member of X^* in the sense of Corollary 2.2.22.

2.2.24 Example. Suppose that $0 \leq p < 1$. It was shown in Examples 2.2.5 and 2.2.6 that the only nonempty convex open subset of $L_p[0,1]$ is $L_p[0,1]$ itself, so $(L_p[0,1])^* = \{0\}$. Therefore $L_p[0,1]$ is a Hausdorff TVS having more than just the zero element in which no two distinct elements can be separated by a member of the dual space in the sense of Corollary 2.2.22.

Let f be a nonzero element of $L_p[0,1]$ and let $Y = \langle\{f\}\rangle$, a one-dimensional subspace of $L_p[0,1]$. Let $y^*(\alpha f) = \alpha$ whenever $\alpha \in \mathbb{F}$. Then y^* is a linear functional on Y that is continuous since its kernel is the closed subspace $\{0\}$ of Y. Since $(L_p[0,1])^* = \{0\}$, there is no member of $(L_p[0,1])^*$ whose restriction to Y is y^*. Also, there is no member of $(L_p[0,1])^*$ whose kernel includes the closed subspace $\{0\}$ of $L_p[0,1]$ and whose value at f is 1. This shows that the assumption of local convexity in Corollaries 2.2.20 and 2.2.21 cannot in general be omitted.

The fact that $(L_0[0,1])^* = \{0\}$ is a result of Otton M. Nikodým [175], while the corresponding property of $L_p[0,1]$ when $0 < p < 1$ is due to Mahlon M. Day [46].

2.2.25 Example. Suppose that $0 < p < 1$. For each m in $\mathbb{N}$, the map x_m^* that sends each member (α_n) of ℓ_p to its m^{th} term α_m is clearly a continuous linear functional on ℓ_p, so ℓ_p^* has enough members to separate the elements of ℓ_p in the sense of Corollary 2.2.22 even though ℓ_p is not an LCS.

If the set F in the statement of Mazur's separation theorem is only required to be nonempty and convex instead of flat, then the conclusion of the theorem turns out to hold in a form that is only slightly weakened.

2.2.26 Eidelheit's Separation Theorem. (M. Eidelheit, 1936 [74]; J. Dieudonné, 1941 [60]). *Let X be a TVS and let C_1 and C_2 be nonempty convex subsets of X such that C_2 has nonempty interior. If $C_1 \cap C_2^\circ = \varnothing$, then there is a member x^* of X^* and a real number s such that*

(1) $\operatorname{Re} x^* x \geq s$ *for each x in C_1;*
(2) $\operatorname{Re} x^* x \leq s$ *for each x in C_2; and*
(3) $\operatorname{Re} x^* x < s$ *for each x in C_2°.*

Proof. Since the flat subset $\{0\}$ of X does not intersect the nonempty convex open subset $C_2^\circ - C_1$ of X, Mazur's separation theorem yields an x^* in X^* such that for each x_2 in C_2° and each x_1 in C_1,

$$\operatorname{Re} x^* x_2 - \operatorname{Re} x^* x_1 = \operatorname{Re} x^*(x_2 - x_1) < \operatorname{Re} x^* 0 = 0.$$

It follows that there is a real number s such that

$$\sup\{\operatorname{Re} x^* x : x \in C_2^\circ\} \leq s \leq \inf\{\operatorname{Re} x^* x : x \in C_1\}.$$

Notice that x^* and s satisfy (1). Now fix an x_2 in C_2° and an x_1 in C_1. The continuity of the vector space operations of X implies that there is a t_0 such that $0 < t_0 < 1$ and $t_0 x_1 + (1 - t_0)x_2 \in C_2^\circ$. Therefore

$$\begin{aligned} s &\geq \operatorname{Re} x^*\big(t_0 x_1 + (1 - t_0)x_2\big) \\ &= t_0 \operatorname{Re} x^* x_1 + (1 - t_0) \operatorname{Re} x^* x_2 \\ &> t_0 \operatorname{Re} x^* x_2 + (1 - t_0) \operatorname{Re} x^* x_2 \\ &= \operatorname{Re} x^* x_2, \end{aligned}$$

from which it follows that x^* and s satisfy (3). Finally, let x be an element of C_2 and let x_2 be the element of C_2° previously fixed. If $0 < t < 1$, then $tx + (1 - t)x_2 \in C_2^\circ$ by Lemma 2.2.18, and so

$$t \operatorname{Re} x^* x + (1 - t) \operatorname{Re} x^* x_2 = \operatorname{Re} x^*\big(tx + (1 - t)x_2\big) < s.$$

Letting t increase to 1 shows that $\operatorname{Re} x^* x \leq s$, which establishes (2) and finishes the proof. ∎

Mazur's and Eidelheit's separation theorems both require one of the convex sets to have nonempty interior. The third and final separation theorem of this section does not, but instead requires one of the convex sets to be closed and the other to be compact, and also requires that the TVS in question be locally convex.

The following lemma gives another example of a phenomenon previously encountered in Proposition 2.2.11: In many ways, a compact subset of a TVS behaves as if it were a singleton.

2.2.27 Lemma. *Suppose that X is a TVS and that A and K are subsets of X such that A is closed and K is compact. Then $A + K$ is closed.*

Proof. Suppose that (a_α) and (k_α) are nets in A and K respectively such that $(a_\alpha + k_\alpha)$ converges to some x in X. It is enough to show that $x \in A + K$. From the compactness of K, there is a subnet (k_β) of (k_α) converging to some k in K. Since $a_\beta = (a_\beta + k_\beta) - k_\beta \to x - k$, it follows that $x - k \in A$, so $x \in k + A \subseteq A + K$. ■

2.2.28 Theorem. (J. W. Tukey, 1942 [233]; V. L. Klee, 1951 [133]). *Let K and C be disjoint nonempty convex subsets of an LCS X such that K is compact and C is closed. Then there is a member x^* of X^* such that* $\max\{ \operatorname{Re} x^*x : x \in K \} < \inf\{ \operatorname{Re} x^*x : x \in C \}$.

Proof. Since $-K$ is compact, it follows from the preceding lemma that $C - K$ is closed, and from the fact that $C \cap K = \varnothing$ that $0 \notin C - K$. By the local convexity of X, there is a convex neighborhood U of 0 such that $U \cap (C - K) = \varnothing$, which implies that $(K + U) \cap C = \varnothing$. Since $K + U$ is a nonempty open convex set disjoint from C, it follows from Eidelheit's separation theorem that there is a member x^* of X^* such that for each x_0 in $K + U$,

$$\operatorname{Re} x^*x_0 < \inf\{ \operatorname{Re} x^*x : x \in C \}.$$

From the compactness of $(\operatorname{Re} x^*)(K)$, there is a k_0 in K such that $\operatorname{Re} x^*k_0 = \sup\{ \operatorname{Re} x^*x : x \in K \}$. Since $k_0 = k_0 + 0 \in K + U$,

$$\max\{ \operatorname{Re} x^*x : x \in K \} = \operatorname{Re} x^*k_0 < \inf\{ \operatorname{Re} x^*x : x \in C \},$$

as required. ■

Tukey proved the preceding theorem for only one special type of locally convex topology, the weak topology of a normed space to be defined later in this chapter. The general case of the theorem is due to Klee. The hypotheses on the two convex sets in the theorem cannot be relaxed very much, as is illustrated by several examples in Tukey's paper and by an example of Dieudonné [61], who showed that ℓ_1 has two disjoint nonempty closed bounded convex subsets that cannot be separated by a bounded linear functional in the sense of the preceding theorem. Klee [134] later extended Dieudonné's result from ℓ_1 to all nonreflexive separable Banach spaces. Also, the hypotheses of the theorem cannot be relaxed by only requiring X to be a TVS; see Exercise 2.32.

One important consequence of Theorem 2.2.28 is that if it is known that the topology of a vector space X is locally convex, and the dual space of X under that topology is known, then the closure of each convex subset of X is completely determined just by that information. This is the content of the following corollary, which will have several major applications later in this chapter.

2.2.29 Corollary. *Suppose that a vector space X has two locally convex topologies $\mathfrak{T}_1$ and $\mathfrak{T}_2$ such that the dual spaces of X under the two topologies are the same. Let C be a convex subset of X. Then the $\mathfrak{T}_1$-closure of C is the same as its $\mathfrak{T}_2$-closure. In particular, the set C is $\mathfrak{T}_1$-closed if and only if it is $\mathfrak{T}_2$-closed.*

PROOF. It may be assumed that $C \neq \varnothing$. Let X^* represent the dual space of X under each of $\mathfrak{T}_1$ and $\mathfrak{T}_2$. For each x in $X \setminus \overline{C}^{\mathfrak{T}_1}$, use Theorem 2.2.28 to produce an x_x^* in X^* such that

$$\operatorname{Re} x_x^* x = \max\{\operatorname{Re} x_x^* y : y \in \{x\}\} < \inf\{\operatorname{Re} x_x^* y : y \in \overline{C}^{\mathfrak{T}_1}\},$$

and let $A_x = \{z : z \in X, \operatorname{Re} x_x^* z \geq \inf\{\operatorname{Re} x_x^* y : y \in \overline{C}^{\mathfrak{T}_1}\}\}$. Then each A_x is $\mathfrak{T}_2$-closed, and $\overline{C}^{\mathfrak{T}_1} = \bigcap\{A_x : x \in X \setminus \overline{C}^{\mathfrak{T}_1}\}$. Therefore $\overline{C}^{\mathfrak{T}_1}$ is $\mathfrak{T}_2$-closed, and similarly $\overline{C}^{\mathfrak{T}_2}$ is $\mathfrak{T}_1$-closed. It follows that $\overline{C}^{\mathfrak{T}_1} = \overline{C}^{\mathfrak{T}_2}$. ■

Suppose that X is a finite-dimensional vector space. Then Corollaries 1.4.17 and 1.4.19 together imply that X has exactly one norm topology and that this topology is induced by a Banach norm. It is even true that this Banach norm topology is the only Hausdorff vector topology that X can have, as will now be shown.

2.2.30 Lemma. *Suppose that X is a Hausdorff TVS and that Y is a subspace of X such that the topology of Y inherited from X is a Banach norm topology. Then Y is a closed subspace of X.*

PROOF. Suppose that (y_α) is a net in Y that converges to an x in X. It is enough to show that $x \in Y$. By Proposition 2.1.47, the net (y_α) is Cauchy with respect to the vector topology of X and therefore with respect to that of Y. It follows from Corollary 2.1.51 that (y_α) converges in Y to some y. Since (y_α) also converges to y in X and X is Hausdorff, it must be that $y = x$, so $x \in Y$ as required. ■

2.2.31 Theorem. *Suppose that X is a finite-dimensional vector space. Then X has exactly one Hausdorff vector topology. This topology is induced by a Banach norm.*

PROOF. The proof is by induction on the dimension of X. Let P_n be the proposition that the conclusions of the theorem hold for X whenever X is a vector space with finite dimension n. It is clear that P_0 is true. Suppose that $n \geq 1$ and that P_{n-1} is true. Let X be a vector space of dimension n and let $\{x_1, \dots, x_n\}$ be a basis for X. Let

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| = |\alpha_1| + \cdots + |\alpha_n|$$

whenever $\alpha_1, \dots, \alpha_n \in \mathbb{F}$. It is easy to check that $\|\cdot\|$ is a norm on X and so is a Banach norm by Corollary 1.4.19. Now let $\mathfrak{T}$ be a Hausdorff vector topology on X and let $X_{\mathfrak{T}}$ and $X_{\|\cdot\|}$ be X equipped with the topology $\mathfrak{T}$ and the topology induced by $\|\cdot\|$ respectively. Let $x_1^*, \ldots, x_n^*$ be the coordinate functionals for $\{x_1, \ldots, x_n\}$; that is, let $x_j^*(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \alpha_j$ when $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ and $j = 1, \ldots, n$. Then the kernel of each of the linear functionals x_j^* has dimension $n-1$ and so is closed in $X_{\mathfrak{T}}$ by P_{n-1} and the lemma preceding this theorem. It follows from Theorem 2.2.16 that each x_j^* is continuous on $X_{\mathfrak{T}}$, and from the defining formula for $\|\cdot\|$ that each x_j^* is continuous on $X_{\|\cdot\|}$. Since $x = (x_1^* x)x_1 + \cdots + (x_n^* x)x_n$ whenever $x \in X$, the continuity of each x_j^* and of the vector space operations of $X_{\mathfrak{T}}$ and $X_{\|\cdot\|}$ assures that the identity operator on X, viewed as a map from $X_{\mathfrak{T}}$ onto $X_{\|\cdot\|}$ or from $X_{\|\cdot\|}$ onto $X_{\mathfrak{T}}$, is continuous. The topologies of $X_{\mathfrak{T}}$ and $X_{\|\cdot\|}$ are therefore the same, which proves P_n and finishes the induction. ∎

2.2.32 Corollary. *Every finite-dimensional subspace of a Hausdorff TVS is a closed subspace of the space.*

PROOF. This follows immediately from the theorem and the lemma preceding it. ∎

2.2.33 Corollary. *Let X be a finite-dimensional Hausdorff TVS and let Y be a TVS. Then every linear operator from X into Y is continuous.*

PROOF. It may be assumed that $X \neq \{0\}$. Let $\{x_1, \ldots, x_n\}$ be a basis for X, and let $\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| = |\alpha_1| + \cdots + |\alpha_n|$ whenever $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. Then the Hausdorff vector topology of X is induced by the norm $\|\cdot\|$. Suppose that T is a linear operator from X into Y. If $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$, then

$$T(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \alpha_1 T x_1 + \cdots + \alpha_n T x_n.$$

It follows from the continuity of the vector space operations of Y and of the maps $\alpha_1 x_1 + \cdots + \alpha_n x_n \mapsto \alpha_j$ when $j = 1, \ldots, n$ that T is continuous. ∎

Exercises

2.18 Let μ be a positive measure on a σ-algebra Σ of subsets of a set Ω.

(a) Prove that the function d in Example 2.2.5 is a metric on $L_0(\Omega, \Sigma, \mu)$.

(b) Prove that a sequence in $L_0(\Omega, \Sigma, \mu)$ is Cauchy with respect to d if and only if the sequence is Cauchy in measure, and that a sequence in $L_0(\Omega, \Sigma, \mu)$ converges to an f in $L_0(\Omega, \Sigma, \mu)$ with respect to d if and only if the sequence converges to f in measure.

(c) Suppose that λ is Lebesgue measure on the σ-algebra Σ of Lebesgue-measurable subsets of $\mathbb{R}$. Find an element f of $L_0(\mathbb{R}, \Sigma, \lambda)$ such that the sequence $(n^{-1} f)$ does not converge in measure to 0. Conclude that multiplication of vectors by scalars is not a continuous operation for $L_0(\mathbb{R}, \Sigma, \lambda)$.

(d) Suppose that $\mu(\Omega) < \infty$. Let

$$\rho(f, g) = \int_\Omega \frac{|f-g|}{1+|f-g|}\, d\mu$$

whenever $f, g \in L_0(\Omega, \Sigma, \mu)$. Prove that ρ is a metric on $L_0(\Omega, \Sigma, \mu)$. Prove that a sequence in $L_0(\Omega, \Sigma, \mu)$ is Cauchy with respect to ρ if and only if the sequence is Cauchy in measure, and that a sequence in $L_0(\Omega, \Sigma, \mu)$ converges to an f in $L_0(\Omega, \Sigma, \mu)$ with respect to ρ if and only if the sequence converges to f in measure. Conclude that ρ is another complete invariant metric that induces the topology of convergence in measure.

2.19 Let μ be the counting measure on the collection Σ of all subsets of $\mathbb{N}$. Then the space ℓ_0 is defined to be $L_0(\mathbb{N}, \Sigma, \mu)$, that is, the space of all sequences of scalars with the metric d of Example 2.2.5.

(a) Prove that

$$d\big((\alpha_n), (\beta_n)\big) = \min\big\{1, \sup\{\, |\alpha_n - \beta_n| : n \in \mathbb{N}\,\}\big\}$$

whenever $(\alpha_n), (\beta_n) \in \ell_0$.

(b) Prove that ℓ_0 is not a TVS.

(c) Prove that the topology that ℓ_∞ inherits as a subspace of ℓ_0 is its usual norm topology.

2.20 Suppose that $0 < p < 1$ and $n \in \mathbb{N}$. Let μ be the counting measure on the σ-algebra Σ of all subsets of $\{1, \ldots, n\}$, and let $\ell_p^n = L_p(\{1, \ldots, n\}, \Sigma, \mu)$. Notice that ℓ_p^n is the space of all ordered n-tuples $(\alpha_1, \ldots, \alpha_n)$ of scalars with the metric given by the formula

$$d\big((\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n)\big) = \sum_{j=1}^{n} |\alpha_j - \beta_j|^p.$$

Prove that ℓ_p^n is an LCS. Now consider the scalar field to be $\mathbb{R}$ and sketch the graph of the closed unit ball of $\ell_{1/2}^2$, that is, the set of points of $\ell_{1/2}^2$ no more than one unit from the origin.

2.21 Suppose that $0 < p < 1$ and that x and y are two different elements of ℓ_p. Show that there are disjoint convex neighborhoods C_x and C_y of x and y respectively. Notice that ℓ_p has many more convex open sets than does $L_p[0,1]$ (which has only two), even though it does not have enough to make it an LCS.

2.22 Suppose that (Ω, Σ, μ) is a finite positive measure space and $0 < p \le \infty$. Prove that the set underlying $L_p(\Omega, \Sigma, \mu)$ is a dense subset of $L_0(\Omega, \Sigma, \mu)$.

2.23 Show that equality need not hold in the inclusions $\overline{A} + \overline{B} \subseteq \overline{A+B}$ and $A^\circ + B^\circ \subseteq (A+B)^\circ$ in Theorem 2.2.9 (d). Also, find a balanced subset of a TVS whose interior is not balanced.

2.24 Prove that if A and B are bounded subsets of a TVS, then $A + B$ is bounded.

2.25 The purpose of this exercise is to show that in a TVS, translation by a compact set has many of the properties of translation by a point. Let A and K be subsets of a TVS such that K is compact. Prove that if A is open, closed, compact, or bounded, then $A+K$ has that same property. (Much of the work has been done for you in various results from this section.)

2.26 Let C be a convex subset of a TVS such that $C^\circ \neq \varnothing$. Prove that C and C° have the same closure.

2.27 Let X and Y be TVSs, let A be a subset of X, and let f be a function from A into Y. Then f is said to be *uniformly continuous* on A if it has this property: For each neighborhood U of 0 in Y there is a neighborhood V_U of 0 in X such that if $x \in A$ and $x' \in A \cap (x + V_U)$, then $f(x') \in f(x) + U$. Prove that if A is compact and f is continuous on A, then f is uniformly continuous on A. (One way to do this is to write out an "epsilon over two" type of proof for the special case in which $X = Y = \mathbb{R}$, then generalize it using Theorem 2.2.9 (g).)

2.28 Give an example of an unbounded convergent net in a TVS.

2.29 Suppose that $0 < p < 1$. If $(\alpha_n) \in \ell_p$ and $(\beta_n) \in \ell_\infty$, let

$$x^*_{(\beta_n)}(\alpha_n) = \sum_n \alpha_n \beta_n.$$

Prove that the map $(\beta_n) \mapsto x^*_{(\beta_n)}$ is a vector space isomorphism from ℓ_∞ onto ℓ_p^*.

2.30 Give an example of a vector space X with a norm $\|\cdot\|$ and a Hausdorff vector topology $\mathfrak{T}$ such that $(X, \|\cdot\|)^*$ and $(X, \mathfrak{T})^*$ do not have the same underlying vector space.

2.31 Obtain Proposition 1.9.15 as a corollary of Eidelheit's separation theorem.

2.32 Find two disjoint nonempty convex subsets K and C of a TVS, with K compact and C closed, such that K and C cannot be separated by a continuous linear functional in the sense of Theorem 2.2.28.

2.33 Let X be a TVS and let S be the collection of all x in X such that each open set containing either x or 0 contains both.

(a) Prove that S is a subspace of X.

(b) Prove that every open set that intersects S actually includes S.

(c) Prove that if X is not completely regular, then $S \neq \{0\}$. Thus, a TVS that is not completely regular has very weak separation properties.

(d) Let π be the quotient map from X onto the quotient vector space X/S; that is, let $\pi(x) = x + S$ for each x in X. Define the *quotient topology* of X/S by letting a subset U of X/S be open if and only if $\pi^{-1}(U)$ is open in X. Show that this defines a completely regular vector topology for X/S. (This is one way to start with a TVS that is not T_0 and construct a completely regular TVS from it.)

(e) Let (Ω, Σ, μ) be a positive measure space and let X be the vector space of all scalar-valued μ-integrable functions on Ω. In this exercise, do *not* consider functions to be the same if they agree almost everywhere but differ somewhere on Ω. For each f in X and each positive ϵ, let $U(f,\epsilon) = \{\, g : g \in X,\ \int_\Omega |g - f| \, d\mu < \epsilon \,\}$. Prove that $\{\, U(f,\epsilon) : f \in X,\ \epsilon > 0 \,\}$ is a basis for a locally convex topology for X. Describe S. Describe X/S and the quotient topology of X/S.

2.34 The purpose of this exercise is to study vector topologies for finite-dimensional vector spaces without assuming that the topologies are Hausdorff. Part (a) of Exercise 2.33 is needed for this. Let $\mathfrak{T}$ be a vector topology for a finite-dimensional vector space X and let S be the subspace of X consisting of all x in X such that each open set containing either x or 0 contains both. Let Y be a vector space complement of S in X, that is, a subspace of X such that $S + Y = X$ and $S \cap Y = \{0\}$. (For example, let $\mathfrak{B}_1$ be a vector space basis for S, let $\mathfrak{B}_2$ be a subset of X such that $\mathfrak{B}_1 \cap \mathfrak{B}_2 = \varnothing$ and $\mathfrak{B}_1 \cup \mathfrak{B}_2$ is a basis for X, and let $Y = \langle \mathfrak{B}_2 \rangle$.)

(a) Prove that the relative topology that Y inherits from X is the unique Hausdorff vector topology for Y.

(b) Prove that $\mathfrak{T} = \{\, U + S : U \text{ is an open subset of } Y \,\}$.

(c) Prove that $\mathfrak{T}$ is a locally convex subtopology of the unique Hausdorff vector topology of X.

This exercise will be continued at the end of Section 2.4 as Exercise 2.51, in which will be developed a characterization of vector topologies on finite-dimensional vector spaces in terms of linear functionals on the space.

*2.3 Metrizable Vector Topologies

None of the material in this section depends on that of any other optional section.

This section contains a few results about metrizable vector topologies that are not, in general, induced by norms. This material is marked as optional since most of the vector topologies important for this book either are induced by norms or are not in general compatible with any metric whatever, as will be seen in Propositions 2.5.14 and 2.6.12. However, it is interesting and instructive to see which of the most important results about normed spaces survive in the somewhat more general setting to be studied in this section.

Since completeness plays such an important rule in the study of normed spaces, it is natural to pay special attention to completeness in the study of more general metrizable TVSs. As was shown in Exercise 1.42, it is quite possible for a vector topology to be induced by two different metrics, one incomplete and the other complete, so it is useful to have a form of completeness that does not depend on the particular metric that induces the vector topology. One possibility would be to use completeness as an abelian topological group under vector addition; see Definition 2.1.48. It is somewhat more customary to use the following form of completeness, and that is what will be done here. As will be shown by Corollary 2.3.22, these two forms of completeness are equivalent for metrizable vector topologies.

2.3.1 Definition. A topological space is *topologically complete* if some complete metric induces its topology.

2.3.2 Definition. An *F-space* is a topologically complete TVS. A *Fréchet space* is a locally convex F-space.

The reader should be warned that some sources include local convexity in the definition of an F-space, while others omit it from that of a Fréchet space. It is also common to require that the topology of a TVS be induced by some complete *invariant* metric before the TVS is called an F-space. Banach noticed that all of the topologically complete TVSs he studied actually do have their topologies induced by complete invariant metrics, and he asked in [13] if this must always be the case. It is a remarkable fact due to Victor Klee that the answer is yes, as will be shown in Theorem 2.3.20. The definition of an F-space that requires invariance is therefore equivalent to the one given above that does not.

2.3.3 Example. Suppose that μ is a finite positive measure on a σ-algebra Σ of subsets of a set Ω. It was shown in Example 2.2.5 that $L_0(\Omega, \Sigma, \mu)$ is a TVS whose topology is induced by a complete metric, so $L_0(\Omega, \Sigma, \mu)$ is an F-space. It was also shown in that same example that $L_0[0, 1]$ is not locally convex, so $L_0[0, 1]$ is an F-space that is not a Fréchet space.

2.3.4 Example. Suppose that $0 < p < 1$. It was shown in Example 2.2.6 that if μ is any positive measure, finite or not, on a σ-algebra Σ of subsets of a set Ω, then $L_p(\Omega, \Sigma, \mu)$ is a TVS whose topology is induced by a complete metric, so $L_p(\Omega, \Sigma, \mu)$ is an F-space. It was also shown in Example 2.2.6 that $L_p[0, 1]$ is not locally convex, and so is an F-space that is not a Fréchet space. The space ℓ_p is a further example of an F-space that is not a Fréchet space, as was seen in Example 2.2.7.

Of course, every Banach space is a locally convex topologically complete TVS. Definition 2.3.2 provides a shorter way to say this.

2.3.5 Theorem. *Every Banach space is a Fréchet space.*

As will be shown by Example 2.3.25, not every Fréchet space has its topology induced by a norm.

This section contains two major results linking metrizability of a vector topology to metrizability by an invariant metric. The first, Theorem 2.3.13, says that a vector topology that is induced by a metric is in fact induced by an invariant metric, while the second, Theorem 2.3.20, says that a vector topology that is induced by a complete metric is actually induced by a complete invariant metric. Before obtaining the first of these two, it is necessary to make a short excursion into the theory of local bases for vector topologies.

2.3.6 Definition. A *local basis* for a vector topology is a basis for the topology at 0.

That is, a local basis for a vector topology is a collection $\mathfrak{B}_0$ of neighborhoods of 0 such that every neighborhood of 0 includes a member of $\mathfrak{B}_0$. Notice that every vector topology has a local basis, namely, the collection of all neighborhoods of 0.

The following proposition contains a few basic facts about local bases that all follow easily from the appropriate definitions and from various parts of Theorem 2.2.9.

2.3.7 Proposition. *Let $\mathfrak{B}_0$ be a local basis for the topology of a TVS X.*

(a) *The collection of all translates of members of $\mathfrak{B}_0$ is a basis for the topology of X.*

(b) *The space X has a local basis $\mathfrak{B}'_0$ for its topology with cardinality no more than that of $\mathfrak{B}_0$ such that every member of $\mathfrak{B}'_0$ is balanced. If X is locally convex, then $\mathfrak{B}'_0$ can be selected so that each of its members is balanced and convex.*

(c) *If U is a neighborhood of 0 in X, then there is a member V of $\mathfrak{B}_0$ such that $V \subseteq \overline{V} \subseteq \overline{V+V} \subseteq U$.*

2.3.8 Corollary. *Every TVS has a local basis for its topology consisting of balanced sets, and every LCS has a local basis for its topology consisting of balanced convex sets.*

2.3.9 Corollary. *A TVS is locally convex if and only if its topology has a local basis consisting of convex sets.*

2.3.10 Corollary. *If a vector space has two vector topologies with a common local basis, then the two topologies are the same.*

If the topology of a TVS is induced by a metric, then the open balls centered at the origin whose radii are reciprocals of positive integers form a countable local basis for the topology. This, together with the following result, shows that in fact the existence of a countable local basis characterizes metrizable vector topologies among all Hausdorff vector topologies.

2.3.11 Theorem. (G. Birkhoff, 1936 [27]; S. Kakutani, 1936 [124]). *Suppose that X is a Hausdorff TVS whose topology has a countable local basis. Then the topology of X is induced by an invariant metric such that the open balls centered at the origin are balanced. If X is locally convex, then its topology is induced by a metric with the preceding properties and for which all open balls are convex.*

PROOF. Let $\mathfrak{T}$ be the given vector topology for X. This proof is based on Construction 2.2.13, so the local basis about to be defined is indexed to match the notation of that construction. As an easy consequence of parts (b) and (c) of Proposition 2.3.7, the topology $\mathfrak{T}$ has a local basis $\{ B(2^{-n}) : n = 0, 1, 2, \dots \}$ such that, for each n, the set $B(2^{-n})$ is balanced (and, if X is an LCS, convex) and $B(2^{-n-1}) + B(2^{-n-1}) \subseteq B(2^{-n})$; call this local basis $\mathfrak{B}_0$. Let $B = B(1)$, then let B and $\{ B(2^{-n}) : n \in \mathbb{N} \}$ be used in Step 1 of Construction 2.2.13 to get the construction started. In that construction, the map $t \mapsto B(t)$ from $\{ 2^{-n} : n = 0, 1, 2, \dots \}$ into $\mathfrak{T}$ is extended to $(0, \infty)$ in such a way that each $B(t)$ is a $\mathfrak{T}$-neighborhood of 0 that is balanced (and convex if each member of $\mathfrak{B}_0$ is convex) and $B(s) \subseteq B(t)$ whenever $0 < s < t$. It is then shown that the formula

$$f(x) = \inf\{ t : t > 0,\, x \in B(t) \}$$

defines a $\mathfrak{T}$-continuous nonnegative-real-valued function on X such that $f(0) = 0$ and such that $f(x) = f(-x)$ and $f(x+y) \le f(x) + f(y)$ whenever $x, y \in X$. If x is a nonzero member of X, then the fact that $\mathfrak{B}_0$ is a local basis for the Hausdorff topology $\mathfrak{T}$ implies that there is a nonnegative integer n such that $x \notin B(2^{-n})$ and therefore that $f(x) \ge 2^{-n} > 0$. It follows from all of this that the formula $d(x, y) = f(x - y)$ defines an invariant metric on X.

For each positive t, let $V(t) = \bigcup\{ B(s) : 0 < s < t \}$. Then each $V(t)$ is a $\mathfrak{T}$-neighborhood of 0 that is balanced (and convex if each member of $\mathfrak{B}_0$ is convex), and $\{ V(t) : t > 0 \}$ is a local basis for $\mathfrak{T}$. It is easy to see that $V(t) = \{ x : x \in X,\, f(x) < t \}$ whenever $t > 0$, that is, that each $V(t)$ is the d-open ball of radius t centered at the origin, so the d-open balls centered at 0 are all balanced (and convex if each member of $\mathfrak{B}_0$ is convex). Now let $U(x, r)$ denote the d-open ball of radius r centered at x. Then for each x in X and each positive r,

$$\begin{aligned} U(x, r) &= \{ y : y \in X,\, f(y - x) < r \} \\ &= x + \{ y : y \in X,\, f(y) < r \} \\ &= x + V(r). \end{aligned}$$

Since $\{ U(x, r) : x \in X,\, r > 0 \}$ is a basis for the topology induced by d and $\{ x + V(r) : x \in X,\, r > 0 \}$ is a basis for $\mathfrak{T}$, and since these bases are the same, the metric d induces $\mathfrak{T}$. Finally, if $\mathfrak{T}$ is a locally convex topology, then the selection of each member of $\mathfrak{B}_0$ to be convex assures that each d-open ball $x + V(r)$ is convex, which finishes the proof. ∎

2.3.12 Corollary. *The topology of a Hausdorff TVS is induced by a metric if and only if the topology has a countable local basis.*

Combining Theorem 2.3.11 with its corollary yields the first of the two major results of this section connecting metrizability of vector topologies to metrizability by invariant metrics.

2.3.13 Theorem. *If the topology of a TVS is induced by a metric, then it is induced by an invariant metric such that the open balls centered at the origin are balanced. If the topology of an LCS is induced by a metric, then it is induced by an invariant metric such that the open balls centered at the origin are balanced and all open balls are convex.*

Notice that the preceding theorem does not say that the *original* metric inducing the topology has to have the nice properties assured by the conclusions of the theorem, but rather that *some* metric inducing the topology has those properties.

The second major result of this section linking metrizability of vector topologies to metrizability by invariant metrics is an easy consequence of a more general result about topological groups obtained by combining the results of three lemmas, each of which is interesting in its own right.

2.3.14 Lemma. *Suppose that X is a group with a topology induced by an invariant metric d_X. Let d_Y be a complete metric on a set Y such that there is an isometry f from X onto a dense subset of Y; that is, let Y be a completion of X. Define a multiplication of members of Y as follows. For each ordered pair (y_1, y_2) of members of Y, let $(x_{1,n})$ and $(x_{2,n})$ be sequences in X such that $\lim_n f(x_{j,n}) = y_j$ when $j = 1, 2$, and let $y_1 \cdot y_2 = \lim_n f(x_{1,n} \cdot x_{2,n})$. Then Y with this multiplication is a group, the map f is a group isomorphism onto a subgroup of Y, and d_Y is invariant.*

PROOF. It must first be shown that multiplication of members of Y is a well-defined operation. Suppose that $y_1, y_2 \in Y$ and that $(x_{1,n})$ and $(x_{2,n})$ are sequences in X such that $\lim_n f(x_{j,n}) = y_j$ when $j = 1, 2$. By the invariance of d_X and the fact that f is an isometry,

$$\begin{aligned} d_Y\big(f(x_{1,n}\cdot x_{2,n}), f(x_{1,m}\cdot x_{2,m})\big) &\le d_X(x_{1,n}\cdot x_{2,n},\, x_{1,m}\cdot x_{2,n}) \\ &\qquad + d_X(x_{1,m}\cdot x_{2,n},\, x_{1,m}\cdot x_{2,m}) \\ &= d_Y\big(f(x_{1,n}), f(x_{1,m})\big) \\ &\qquad + d_Y\big(f(x_{2,n}), f(x_{2,m})\big) \end{aligned}$$

whenever $m, n \in \mathbb{N}$. Since $\big(f(x_{1,n})\big)$ and $\big(f(x_{2,n})\big)$ are Cauchy, the sequence $\big(f(x_{1,n}\cdot x_{2,n})\big)$ is also Cauchy, and therefore converges. Now suppose that $(w_{1,n})$ and $(w_{2,n})$ are any sequences in X such that $\lim_n f(w_{j,n}) = y_j$ when $j = 1, 2$. Letting $(v_{j,n})$ be the sequence $(x_{j,1}, w_{j,1}, x_{j,2}, w_{j,2}, \ldots)$ when

$j = 1, 2$ and observing that $(f(v_{1,n} \cdot v_{2,n}))$ must converge shows that there is no ambiguity in the definition of $y_1 \cdot y_2$. Multiplication is therefore a well-defined operation on Y.

Suppose that $y_1, y_2, y_3 \in Y$ and that sequences $(x_{1,n})$, $(x_{2,n})$, and $(x_{3,n})$ in X are such that $\lim_n f(x_{j,n}) = y_j$ when $j = 1, 2, 3$. Then

$$\begin{aligned}(y_1 \cdot y_2) \cdot y_3 &= \lim_n f\big((x_{1,n} \cdot x_{2,n}) \cdot x_{3,n}\big) \\ &= \lim_n f\big(x_{1,n} \cdot (x_{2,n} \cdot x_{3,n})\big) \\ &= y_1 \cdot (y_2 \cdot y_3),\end{aligned}$$

so multiplication of elements of Y is associative. Let e be the identity of X and let $e_n = e$ for each positive integer n. Then

$$y_1 \cdot f(e) = \lim_n f(x_{1,n} \cdot e_n) = \lim_n f(x_{1,n}) = y_1,$$

and similarly $f(e) \cdot y_1 = y_1$, so $f(e)$ is an identity for Y. Notice next that the invariance of d_X implies that $d_X(x_{1,n}^{-1}, x_{1,m}^{-1}) = d_X(x_{1,m}, x_{1,n})$ whenever $m, n \in \mathbb{N}$, so $\big(f(x_{1,n}^{-1})\big)$ is a Cauchy sequence in Y. Let y_0 denote its limit. Then

$$y_1 \cdot y_0 = \lim_n f(x_{1,n} \cdot x_{1,n}^{-1}) = f(e) = \lim_n f(x_{1,n}^{-1} \cdot x_{1,n}) = y_0 \cdot y_1,$$

so each member of Y has an inverse in Y. It follows that $(Y, \cdot)$ is a group. Also,

$$\begin{aligned}d_Y(y_1 \cdot y_2,\, y_1 \cdot y_3) &= \lim_n d_Y\big(f(x_{1,n} \cdot x_{2,n}), f(x_{1,n} \cdot x_{3,n})\big) \\ &= \lim_n d_X(x_{1,n} \cdot x_{2,n},\, x_{1,n} \cdot x_{3,n}) \\ &= \lim_n d_X(x_{2,n}, x_{3,n}) \\ &= \lim_n d_Y\big(f(x_{2,n}), f(x_{3,n})\big) \\ &= d_Y(y_2, y_3),\end{aligned}$$

and similarly $d_Y(y_2 \cdot y_1,\, y_3 \cdot y_1) = d_Y(y_2, y_3)$, so d_Y is invariant. Finally, if $u, v \in X$, and $u_n = u$ and $v_n = v$ for each positive integer n, then

$$f(u \cdot v) = \lim_n f(u_n \cdot v_n) = f(u) \cdot f(v),$$

which together with the fact that f is one-to-one shows that f is a group isomorphism onto a subgroup of Y. ■

Recall that a G_δ subset of a topological space is a set that is the intersection of countably many open sets.

2.3.15 Lemma. (W. Sierpiński, 1928 [213]). *Suppose that X is a topologically complete subset of a metric space Y. Then X is a G_δ subset of Y.*

PROOF. Let d_Y be the metric of Y and let d_X be a complete metric for X that induces the same topology for X as does d_Y. For each x in X and each n in $\mathbb{N}$, let $r_n(x)$ be a positive real number such that $r_n(x) < n^{-1}$ and $d_X(w,x) < n^{-1}$ whenever $w \in X$ and $d_Y(w,x) < r_n(x)$, and let $U_n(x)$ be the open ball in Y of d_Y-radius $r_n(x)$ centered at x. Let $G_n = \bigcup\{\, U_n(x) : x \in X \,\}$ for each n in $\mathbb{N}$, and let $\Gamma = \bigcap\{\, G_n : n \in \mathbb{N} \,\}$. Then Γ is a G_δ subset of Y. Since each x in X lies in $U_n(x)$ for every n and therefore lies in each G_n, it follows that $X \subseteq \Gamma$. The lemma will be proved once it is shown that $\Gamma \subseteq X$.

Let x_0 be a member of Γ. For each positive integer n, the fact that $x_0 \in G_n$ implies that there is an x_n in X such that $x_0 \in U_n(x_n)$, that is, such that

$$d_Y(x_0, x_n) < r_n(x_n) < \frac{1}{n}. \tag{2.1}$$

It follows that $x_n \to x_0$ in Y.

Let ϵ be a positive number, and choose a positive integer N such that $2/N < \epsilon$. Let m be a positive integer such that

$$\frac{1}{m} < r_N(x_N) - d_Y(x_0, x_N). \tag{2.2}$$

It follows from (2.1) that for each positive integer k,

$$d_Y(x_k, x_N) \le d_Y(x_k, x_0) + d_Y(x_0, x_N) < \frac{1}{k} + d_Y(x_0, x_N),$$

which together with (2.2) implies that if $k > m$, then $d_Y(x_k, x_N) < r_N(x_N)$ and therefore $d_X(x_k, x_N) < N^{-1}$. If $k, l > m$, then

$$d_X(x_k, x_l) \le d_X(x_k, x_N) + d_X(x_N, x_l) < \frac{2}{N} < \epsilon,$$

and therefore the sequence (x_n) is d_X-Cauchy and so convergent to some member of X. Since $x_n \to x_0$ in Y, it follows that $x_0 \in X$, so $\Gamma \subseteq X$. ■

The next lemma is not one of the three important lemmas used directly in the proof of Theorem 2.3.18, but is instead just a "sublemma" for Lemma 2.3.17. The result is almost obvious, but needs a moment's arguing.

2.3.16 Lemma. *Suppose that Y is a topological group, that A is a subset of Y that is of the first category in Y, and that $y \in Y$. Then $y \cdot A$ is of the first category in Y.*

PROOF. Suppose that N is a nowhere dense subset of Y. By an obvious argument involving the definition of a first category set, it suffices to show that $y \cdot N$ is nowhere dense in Y. It follows from the continuity of the group operation that $y \cdot \overline{N} = \overline{y \cdot N}$; one straightforward argument showing this uses Proposition 2.1.18. If $y \cdot N$ were not nowhere dense in Y, then there would be a nonempty open set U in $\overline{y \cdot N}$, so $y^{-1} \cdot U$ would be a nonempty open subset of $\overline{N}$, which would contradict the fact that N is nowhere dense in Y. ■

The following lemma is due to Victor Klee, and the proof given here is from Klee's paper. As Klee points out, the argument is essentially the same as one used by S. Mazur and L. Sternbach [165] to show that every G_δ subspace of a Banach space is closed.

2.3.17 Lemma. (V. L. Klee, 1952 [135]). *Suppose that Y is a topological group that is of the second category in itself and that X is a subgroup of Y that is a dense G_δ subset of Y. Then $X = Y$.*

PROOF. It follows from Proposition 1.5.3 (c) that $Y \setminus X$ is of the first category in Y and therefore that X is of the second category in Y. If $y \in Y \setminus X$, then $y \cdot X \subseteq Y \setminus X$, and so $y \cdot X$ is of the first category in Y, which implies that X itself is of the first category in Y, a contradiction. It follows that $Y \setminus X = \varnothing$ and therefore that $X = Y$. ■

Notice that by Proposition 2.1.43, the hypotheses of the following theorem do imply that the group in question is a topological group.

2.3.18 Theorem. (V. L. Klee, 1952 [135]). *Suppose that X is a group with a topologically complete topology induced by an invariant metric d_X. Then d_X is a complete metric.*

PROOF. By Lemma 2.3.14, there is a group Y with a topology induced by a complete invariant metric d_Y and a map $f\colon X \to Y$ that is an isometry and a group isomorphism onto a dense subgroup of Y. It follows from Proposition 2.1.43 that the group Y with the topology induced by d_Y is a topological group. If ρ_X is a complete metric that induces the topology of X, then it is easy to check that the formula $\rho_Y(f(x_1), f(x_2)) = \rho_X(x_1, x_2)$ defines a complete metric on $f(X)$ that induces the topology that $f(X)$ inherits from Y. As a topologically complete subset of Y, the set $f(X)$ is a G_δ subset of Y by Lemma 2.3.15. The Baire category theorem assures that Y is of the second category in itself, so an application of Lemma 2.3.17 shows that $f(X) = Y$ and therefore that d_X is complete. ■

In [135], Klee used an example due to Dieudonné [62] to show that the conclusion of the preceding theorem can fail if the metric d_X is only assumed to be left-invariant.

As an easy consequence of Proposition 1.4.14 (c), a normed space is a Banach space if there is some complete norm that induces its topology, and Corollaries 2.1.50 and 2.1.51 provide a strengthening of that statement: A normed space is a Banach space if there is some complete invariant metric that induces its topology. It is an immediate consequence of Theorem 2.3.18 that a further strengthening can be made: A normed space is a Banach space if there is some complete metric, invariant or not, that induces its topology.

2.3.19 Corollary. (V. L. Klee, 1952 [135]). *A normed space is a Banach space if and only if it is topologically complete.*

The main consequence of Theorem 2.3.18 for the purposes of this section is the following one, which gives an affirmative answer to Banach's question mentioned after Definition 2.3.2.

2.3.20 Theorem. (V. L. Klee, 1952 [135]). *A TVS is an F-space if and only if its topology is induced by a complete invariant metric.*

PROOF. A TVS whose topology is induced by a complete invariant metric is obviously an F-space. Conversely, suppose that X is an F-space. By Theorem 2.3.13, the topology of X is induced by an invariant metric, and Theorem 2.3.18 assures that this metric is complete. ■

A bit more can be said. The invariant metric used in the proof of the preceding theorem can, by Theorem 2.3.13, be selected to have some special properties. This is summarized by the following result.

2.3.21 Corollary. *Every F-space has its topology induced by a complete invariant metric such that the open balls centered at the origin are balanced. Every Fréchet space has its topology induced by a complete invariant metric such that the open balls centered at the origin are balanced and all open balls are convex.*

Combining Theorems 2.3.13 and 2.3.20 with Corollary 2.1.50 yields the following result immediately.

2.3.22 Corollary. *A TVS whose topology is induced by a metric is an F-space if and only if it is complete as an abelian topological group under vector addition.*

That is, for a metrizable TVS, topological completeness is equivalent to completeness as an abelian topological group under vector addition.

Since the boundedness of certain sets is an issue in many of the remaining results of this section, this is a good place to mention again a convention adopted after Definition 2.2.8. *When it is said that a subset of a TVS is bounded, it is meant that the set is bounded in the topological sense of Definition 2.2.8 rather than in any metric sense, unless specifically stated* *otherwise.* This is especially important when working with metrizable vector topologies, for such topologies are always compatible with a metric for which the entire space is metrically bounded; see Exercise 2.36. However, also keep in mind that for subsets of a normed space, metric boundedness and boundedness as a subset of the corresponding TVS are equivalent.

2.3.23 Definition. A vector topology is *locally bounded* if some neighborhood of the origin in the space is bounded.

For example, every norm topology is locally bounded, since the open unit ball of the space is bounded. The following result is a partial converse of that fact.

2.3.24 Theorem. *Every locally bounded Hausdorff vector topology is induced by a metric.*

PROOF. Suppose that a Hausdorff TVS X has a bounded neighborhood V of 0. It follows that if U is a neighborhood of 0, then there is a positive integer n_U such that $n_U^{-1}V \subseteq U$, so $\{ n^{-1}V : n \in \mathbb{N} \}$ is a countable local basis for the topology of X. By Theorem 2.3.11, the topology of X is metrizable. ■

There are, however, metrizable vector topologies that are not locally bounded.

2.3.25 Example. Let X be the collection of all sequences of scalars, made into a vector space with the usual vector space operations for spaces of sequences. Define $d\colon X \times X \to [0,1]$ by the formula

$$d(x,y) = \sum_j 2^{-j} \min\{1, |x_j - y_j|\}.$$

It is easy to check that d is an invariant metric and that a sequence $(x^{(n)})$ of members of X converges to some x if and only if $\lim_n x_j^{(n)} = x_j$ for each j, for which reason the topology induced by d is called the *topology of termwise convergence.* By Proposition 2.1.43, the invariance of d implies that addition of vectors is continuous. Also, if a sequence $(x^{(n)})$ of members of X converges to some x and $\alpha_n \to \alpha$ in $\mathbb{F}$, then $\lim_n \alpha_n x_j^{(n)} = \alpha x_j$ for each j, and so $\lim_n \alpha_n x^{(n)} = \alpha x$, from which it follows that multiplication of vectors by scalars is continuous. The topology of termwise convergence is therefore a vector topology for X. Since every Cauchy sequence of members of X is termwise Cauchy, hence termwise convergent and therefore convergent to some member of X, the metric d is complete. The space X with the topology of termwise convergence is therefore an F-space. Now define a collection of neighborhoods of the origin of X by letting

$$V_n = \{ x : x \in X,\ |x_j| < n^{-1} \text{ when } j = 1, \ldots, n \}$$

for each positive integer n. It is easy to check that each V_n is convex and that each open ball centered at 0 includes some V_n, from which it follows that $\{ V_n : n \in \mathbb{N} \}$ is a local basis for the topology of X consisting of convex sets and thus that X with the topology of termwise convergence is a Fréchet space.

It will now be shown that the topology of X is not locally bounded, for which it is sufficient to show that no member of $\{ V_n : n \in \mathbb{N} \}$ is bounded. Fix a positive integer n and a positive real number t and let x be the member of X for which $x_{n+1} = t$ and all other terms are 0. Then $x \in V_n \setminus tV_{n+1}$, which shows that V_n is not included in any positive scalar multiple of V_{n+1} and therefore is not bounded.

Notice that since no neighborhood of the origin of X is bounded, the topology of X is not induced by a norm, so X is an example of a Fréchet space whose topology is not compatible with any norm.

In light of the preceding results on metrizability, it is natural to ask what conditions on a TVS assure normability. The following theorem completely settles the question.

2.3.26 Theorem. *A topology for a vector space is induced by a norm if and only if it is a Hausdorff vector topology that is locally bounded and locally convex.*

PROOF. The topology induced by a norm on a vector space is a Hausdorff locally convex topology for which the open unit ball is a bounded neighborhood of the origin. Conversely, suppose that the topology $\mathfrak{T}$ of a Hausdorff TVS X is locally bounded and locally convex. An application of Corollary 2.3.8 produces a neighborhood V of 0 that is bounded, balanced, and convex (and, by Theorem 2.2.9 (f), absorbing). By Proposition 1.9.14, the Minkowski functional p of V is a seminorm on X. If x is a nonzero member of X, then the fact that $\mathfrak{T}$ is Hausdorff implies that $X \setminus \{x\}$ is a neighborhood of 0, which together with the boundedness of V implies that there is a positive t such that $sV \subseteq X \setminus \{x\}$ whenever $0 < s < t$, which in turn implies that $p(x) > 0$. It follows that p is actually a norm. By Proposition 1.9.14,

$$\{ x : x \in X, \, p(x) < 1 \} \subseteq V \subseteq \{ x : x \in X, \, p(x) \leq 1 \}.$$

If $x \in V$, then it follows from the $\mathfrak{T}$-continuity of multiplication of vectors by scalars and the fact that V is $\mathfrak{T}$-open that there is some real number r greater than 1 such that $rx \in V$, which in turn implies that $p(x) < 1$. Therefore V is the open unit ball for p, so the collection $\{ n^{-1}V : n \in \mathbb{N} \}$ is a local basis for the p-topology of X. A glance at the proof of Theorem 2.3.24 shows that this collection is also a local basis for $\mathfrak{T}$, so $\mathfrak{T}$ is induced by the norm p. ∎

As is true for the theory of linear operators between normed spaces, the theory of linear operators between TVSs is quite rich, especially when the topology of the domain space is metrizable. The following definition and theorem extend Definition 1.4.1 and Theorem 1.4.2 from normed spaces into these more general settings.

2.3.27 Definition. Let X and Y be TVSs. A linear operator T from X into Y is *bounded* if $T(B)$ is a bounded subset of Y whenever B is a bounded subset of X. The collection of all bounded linear operators from X into Y is denoted by $B(X,Y)$, or by just $B(X)$ if $X = Y$.

2.3.28 Theorem. *Let X and Y be TVSs and let $T\colon X \to Y$ be a linear operator. Then the following two statements are equivalent.*

(a) *The operator T is continuous.*
(b) *The operator T is continuous at 0.*

Each of the above two statements implies the following one, and all three statements are equivalent if the topology of X is metrizable.

(c) *The operator T is bounded.*

PROOF. Suppose first that T is continuous at 0. If a net (x_α) in X converges to some x, then $x_\alpha - x \to 0$, so $T(x_\alpha - x) \to T0 = 0$, and therefore $Tx_\alpha \to Tx$. The operator T is thus continuous at each point of X, and so (b) $\Rightarrow$ (a). The reverse implication is obvious, so (a) $\Leftrightarrow$ (b).

Now suppose that T is continuous, that B is a bounded subset of X, and that V is a neighborhood of 0 in Y. The continuity of T implies that there is a neighborhood U of 0 in X such that $T(U) \subseteq V$, and the boundedness of B implies that there is a positive s such that $B \subseteq tU$ whenever $t > s$. It follows that $T(B) \subseteq tT(U) \subseteq tV$ whenever $t > s$, and therefore that $T(B)$ is bounded. Statement (a) therefore implies (c).

Finally, suppose that the topology of X is induced by a metric, that T is bounded, and that (x_n) is a sequence in X converging to 0. For each positive integer k, there is a positive integer n_k such that kx_n lies in the open ball of radius k^{-1} centered at 0 whenever $n \geq n_k$, from which it follows that there is a nondecreasing sequence (k_n) of positive integers such that $k_n \to \infty$ and $k_n x_n \to 0$. Since the set $\{\, k_n x_n : n \in \mathbb{N} \,\}$ and the operator T are bounded, so is the set $\{\, k_n T x_n : n \in \mathbb{N} \,\}$. Let W be a neighborhood of 0 in Y and let s be a positive number such that $\{\, k_n T x_n : n \in \mathbb{N} \,\} \subseteq tW$ whenever $t > s$. It follows that there is a positive integer n_W such that $k_n T x_n \in k_n W$ whenever $n \geq n_W$, which implies that $Tx_n \in W$ for all sufficiently large n. The sequence (Tx_n) therefore converges to 0, which establishes the continuity of T at 0 and proves that (c) $\Rightarrow$ (b) when the topology of X is induced by a metric. ■

It is not always true that part (c) of the preceding theorem implies parts (a) and (b) when the topology of X is not metrizable. See Exercise 2.57.

The open mapping theorem, closed graph theorem, and uniform boundedness principle all have natural extensions to F-spaces that will now be obtained. To make the statement of each extension conform as closely as possible to the statement of the corresponding result for normed spaces given in Section 1.6, reference will be made in the statement of each to bounded linear operators. In each case, the domains of the linear operators will be F-spaces, so for these operators boundedness will be equivalent to continuity.

Several of the following results cite Banach's 1932 monograph [13] in addition to earlier sources. When this occurs, the earlier references are to proofs for Banach spaces, while the references to [13] are to Banach's extensions of the results to F-spaces by substantially the same arguments.

2.3.29 The Open Mapping Theorem for F-Spaces. (J. Schauder, 1930 [208]; S. Banach, 1932 [13]). *Every bounded linear operator from an F-space onto an F-space is an open mapping.*

PROOF. This proof uses the fact that if d is an invariant metric on a vector space W and $w_1, \ldots, w_n \in W$, then $d(\sum_{j=1}^n w_j, 0) \leq \sum_{j=1}^n d(w_j, 0)$. This follows from a straightforward induction argument that begins with the observation that $d(w_1 + w_2, 0) = d(w_1, -w_2) \leq d(w_1, 0) + d(0, -w_2) = d(w_1, 0) + d(w_2, 0)$.

Let T be a bounded linear operator from an F-space X onto an F-space Y and let N be a neighborhood of the origin 0_X of X. Suppose that it were shown that $T(N)$ must include a neighborhood of the origin 0_Y of Y. It would follow that if G is an open subset of X and $x \in G$, then

$$T(G) = Tx + T(-x + G) \supseteq Tx + \big(T(-x + G)\big)^\circ,$$

and so $T(G)$ would be open since it would include a neighborhood of each of its points. It is therefore enough to prove that $0_Y \in \big(T(N)\big)^\circ$.

It will first be shown that $0_Y \in \left(\overline{T(N)}\right)^\circ$. Let V be a balanced neighborhood of 0_X such that $V + V \subseteq N$. If $\left(\overline{T(V)}\right)^\circ \neq \varnothing$, then

$$0_Y \in \left(\overline{T(V)}\right)^\circ - \left(\overline{T(V)}\right)^\circ \subseteq \overline{T(V)} - \overline{T(V)} = \overline{T(V)} + \overline{T(V)} \subseteq \overline{T(N)};$$

that is, the set $\overline{T(N)}$ includes the neighborhood $\left(\overline{T(V)}\right)^\circ - \left(\overline{T(V)}\right)^\circ$ of 0_Y. It will therefore follow that $0_Y \in \left(\overline{T(N)}\right)^\circ$ once it is shown that $\left(\overline{T(V)}\right)^\circ$ is not empty. Since $T(X) = Y$ and V is absorbing, it follows that $Y = \bigcup_n T(nV)$, so by the Baire category theorem there must be a positive integer n_0 such that $T(n_0 V)$ is not nowhere dense in Y, that is, such that $\left(\overline{T(n_0 V)}\right)^\circ \neq \varnothing$. It follows that $\left(\overline{T(V)}\right)^\circ \neq \varnothing$ and therefore that $0_Y \in \left(\overline{T(N)}\right)^\circ$.

Let d_X and d_Y be complete invariant metrics inducing the topologies of X and Y respectively. Let $U_X(r)$ and $U_Y(r)$ denote the open balls of radius r centered at 0_X and 0_Y respectively when $r > 0$, and let ϵ be a positive number such that $U_X(\epsilon) \subseteq N$. By the argument of the preceding paragraph, there is a sequence (δ_n) of positive reals converging to 0 such that $U_Y(\delta_n) \subseteq \overline{T(U_X(2^{-n}\epsilon))}$ whenever $n \in \mathbb{N}$. Let y_0 be an arbitrary element of $U_Y(\delta_1)$. The theorem will be proved once it is shown that there is an x_0 in $U_X(\epsilon)$ such that $Tx_0 = y_0$.

Since $y_0 \in U_Y(\delta_1) \subseteq \overline{T(U_X(2^{-1}\epsilon))}$, there is an x_1 in $U_X(2^{-1}\epsilon)$ such that $d_Y(y_0, Tx_1) < \delta_2$. Since $y_0 - Tx_1 \in U_Y(\delta_2) \subseteq \overline{T(U_X(2^{-2}\epsilon))}$, there is an x_2 in $U_X(2^{-2}\epsilon)$ such that $d_Y(y_0, Tx_1 + Tx_2) = d_Y(y_0 - Tx_1, Tx_2) < \delta_3$. Continuing in the obvious fashion yields a sequence (x_n) in X such that $x_n \in U_X(2^{-n}\epsilon)$ and $d_Y(y_0, \sum_{j=1}^{n} Tx_j) < \delta_{n+1}$ for each positive integer n. If $m_1, m_2 \in \mathbb{N}$ and $m_1 < m_2$, then

$$\begin{aligned} d_X\Big(\sum_{j=1}^{m_2} x_j, \sum_{j=1}^{m_1} x_j\Big) &= d_X\Big(\sum_{j=m_1+1}^{m_2} x_j, 0_X\Big) \\ &\le \sum_{j=m_1+1}^{m_2} d_X(x_j, 0_X) \\ &< \sum_{j=m_1+1}^{\infty} 2^{-j}\epsilon \\ &= 2^{-m_1}\epsilon, \end{aligned}$$

from which it follows that the partial sums of the formal series $\sum_n x_n$ form a Cauchy sequence and therefore that $\sum_n x_n$ converges. Let $x_0 = \sum_n x_n$. Since $\lim_n d_Y\big(y_0, T(\sum_{j=1}^{n} x_j)\big) = 0$, it follows that

$$Tx_0 = T\Big(\lim_n \sum_{j=1}^{n} x_j\Big) = \lim_n T\Big(\sum_{j=1}^{n} x_j\Big) = y_0.$$

Finally,

$$d_X(x_0, 0_X) = \lim_n d_X\Big(\sum_{j=1}^{n} x_j, 0_X\Big) \le \sum_{j=1}^{\infty} d_X(x_j, 0_X) < \sum_{j=1}^{\infty} 2^{-j}\epsilon = \epsilon,$$

so $x_0 \in U_X(\epsilon)$ as required. ■

2.3.30 Corollary. (S. Banach, 1929 [11], 1932 [13]). *Every one-to-one bounded linear operator from an F-space onto an F-space has a bounded inverse.*

The preceding corollary is a generalization of Corollary 1.6.6 and, like that earlier result, is sometimes called the inverse mapping theorem.

2.3.31 The Closed Graph Theorem for F-Spaces. (S. Banach, 1932 [13]). *Let T be a linear operator from an F-space X into an F-space Y. Suppose that whenever a sequence (x_n) in X converges to some x in X and (Tx_n) converges to some y in Y, it follows that $y = Tx$. Then T is bounded.*

PROOF. Let d_X and d_Y be complete invariant metrics inducing the topologies of X and Y respectively. For each pair of elements (x_1, y_1) and (x_2, y_2) of $X \times Y$, let

$$d_{X\times Y}\big((x_1,y_1),(x_2,y_2)\big) = \Big(\big(d_X(x_1,x_2)\big)^2 + \big(d_Y(y_1,y_2)\big)^2\Big)^{1/2}.$$

Then $d_{X\times Y}$ is a complete invariant metric that induces the product topology of $X \times Y$; see the discussion of product metrics in Appendix B. It is easy to check that $X \times Y$ is an F-space when given its product topology and the usual vector space operations for a vector space sum. Let $G = \{\,(x, Tx) : x \in X\,\}$; that is, let G be the graph of T in $X \times Y$. It follows from the hypotheses of this theorem that G is a closed subspace of $X \times Y$ and so is itself an F-space. Since the map $(x, Tx) \mapsto x$ from G onto X is a one-to-one bounded linear operator, its inverse is bounded by Corollary 2.3.30, so the map $x \mapsto (x, Tx) \mapsto Tx$ is itself a bounded linear operator. ■

The statement of the uniform boundedness principle given in Section 1.6 needs a bit of reinterpretation before it can be extended from Banach spaces to F-spaces. Let $\mathfrak{F}$ be a nonempty family of bounded linear operators from a Banach space X into a normed space Y. Suppose that $\sup\{\,\|Tx\| : T \in \mathfrak{F}\,\}$ is finite for each x in X, which is the same as saying that $\{\,Tx : T \in \mathfrak{F}\,\}$ is bounded whenever $x \in X$. The conclusion of the uniform boundedness principle for Banach spaces is that $\sup\{\,\|T\| : T \in \mathfrak{F}\,\}$ is finite. A moment's thought about properties of the norm of a bounded linear operator between normed spaces shows that this conclusion is equivalent to the following statement: For each bounded subset A of X, there is a bounded subset B_A of Y such that $T(A) \subseteq B_A$ whenever $T \in \mathfrak{F}$.

2.3.32 Definition. A family $\mathfrak{F}$ of linear operators from a TVS X into a TVS Y is *uniformly bounded* if $\bigcup\{\,T(B) : T \in \mathfrak{F}\,\}$ is a bounded subset of Y whenever B is a bounded subset of X.

In addition to the horde of citations that accompanies the statement of the uniform boundedness principle for Banach spaces, the following result contains one more, since Mazur and Orlicz obtained the extension to F-spaces.

2.3.33 The Uniform Boundedness Principle for F-Spaces. (H. Hahn, 1922 [98]; S. Banach, 1922 [10]; T. H. Hildebrandt, 1923 [104]; S. Banach and H. Steinhaus, 1927 [17]; S. Mazur and W. Orlicz, 1933 [164]). *Let $\mathfrak{F}$ be a family of bounded linear operators from an F-space X into a TVS Y. Suppose that $\{Tx : T \in \mathfrak{F}\}$ is bounded for each x in X. Then $\mathfrak{F}$ is uniformly bounded. In short, the pointwise boundedness of $\mathfrak{F}$ implies its uniform boundedness.*

PROOF. To avoid having to think about a special case throughout this proof, notice that it can be assumed that $\mathfrak{F} \neq \varnothing$. Let B be a bounded subset of X and U a neighborhood of the origin 0_Y of Y. The theorem will be proved once a positive s is found such that $T(B) \subseteq tU$ when $T \in \mathfrak{F}$ and $t > s$. Let V be a balanced neighborhood of 0_Y such that $\overline{V+V} \subseteq U$ and let $S = \bigcap\{T^{-1}(\overline{V}) : T \in \mathfrak{F}\}$, a closed subset of X because of the continuity of each T in $\mathfrak{F}$. If $x \in X$, then the boundedness of $\{Tx : T \in \mathfrak{F}\}$ assures that $\{Tx : T \in \mathfrak{F}\} \subseteq n_x V$ for some positive integer n_x and therefore that $x \in n_x S$. It follows that $X = \bigcup\{nS : n \in \mathbb{N}\}$. By the Baire category theorem, one of the closed sets nS, and hence S itself, must have nonempty interior. Let x_0 be a point in S° and let $W = x_0 - S^\circ$, a neighborhood of the origin of X. For each T in $\mathfrak{F}$,

$$T(W) \subseteq Tx_0 - T(S) \subseteq \overline{V} - \overline{V} \subseteq \overline{V+V} \subseteq U.$$

The boundedness of B yields a positive s such that $B \subseteq tW$ whenever $t > s$. It follows that if $T \in \mathfrak{F}$ and $t > s$, then $T(B) \subseteq tT(W) \subseteq tU$, as required. ■

In the following corollary, the reason for requiring Y to have a Hausdorff topology is to assure that limits of convergent sequences in Y are unique.

2.3.34 Corollary. *Let (T_n) be a sequence of bounded linear operators from an F-space X into a Hausdorff TVS Y such that $\lim_n T_n x$ exists for each x in X. Define $T\colon X \to Y$ by the formula $Tx = \lim_n T_n x$. Then T is a bounded linear operator from X into Y.*

PROOF. The continuity of the vector space operations of Y and the linearity of each T_n together imply the linearity of T. Let B be a bounded subset of X. The proof will be complete once it is shown that $T(B)$ is bounded. The convergence of the sequence $(T_n x)$ for each x in X forces $\{T_n x : n \in \mathbb{N}\}$ to be bounded whenever $x \in X$, which by the preceding theorem implies that $\bigcup\{T_n(B) : n \in \mathbb{N}\}$ is bounded. As a subset of the bounded set $\overline{\bigcup\{T_n(B) : n \in \mathbb{N}\}}$, the set $T(B)$ is bounded. ■

Thus, the three fundamental theorems of Section 1.6, as well as their major corollaries, survive virtually unscathed when extended from Banach spaces to F-spaces. This leads to the question of which major results about normed spaces do not generalize to metrizable TVSs.

Some of the most important casualties are the normed space version of the Hahn-Banach extension theorem and its main corollaries about continuous linear functionals. Most of these results do extend in some way to Fréchet spaces, as is shown by Corollaries 2.2.20–2.2.22 of Mazur's separation theorem, but this is due to the local convexity of Fréchet spaces (and, in the case of Corollary 2.2.22, to the fact that their topologies are Hausdorff) rather than to the metrizability of their topologies. These corollaries do not in general extend to F-spaces that are not locally convex. For instance, it was shown in Examples 2.2.24 and 2.2.25 that if $0 < p < 1$, then Corollary 2.2.22, which says that Hausdorff LCSs have enough continuous linear functionals to separate points, does extend to the non-locally-convex F-space ℓ_p but not to $L_p[0,1]$.

The situation is less ambiguous for Corollary 2.2.21, which says that every LCS has the property that for each subspace of the space, every continuous linear functional on that subspace has an extension as a continuous linear functional to the entire space. For F-spaces, this property is called the *Hahn-Banach extension property.* In 1969, Duren, Romberg, and Shields [69] noted that all of the classical non-locally-convex F-spaces, as well as other examples of such spaces that they had studied, lack the Hahn-Banach extension property, and asked if it might be the case that an F-space has the Hahn-Banach extension property if *and only if* it is locally convex. Nigel Kalton showed in 1974 [126] that the answer is yes. For an exposition of Kalton's result, see [128], which is an excellent source for the reader wishing to learn more about F-spaces.

Exercises

2.35 Let X be the vector space of all continuous functions from $\mathbb{F}$ into $\mathbb{F}$, and let

$$d(f,g) = \sum_j 2^{-j} \min\{1, \max\{\, |f(\alpha) - g(\alpha)| : |\alpha| \le j\,\}\}$$

whenever $f, g \in X$.

(a) Prove that d is a complete invariant metric.

(b) Prove that a sequence (f_n) in X converges to an f in X if and only if the following holds: For each compact subset K of $\mathbb{F}$, the sequence (f_n) converges uniformly to f on K. (Because of this, the topology induced by d is called the *topology of uniform convergence on compact sets.*)

(c) Prove that X with its d-topology is a Fréchet space that is not locally bounded. Conclude that this topology is not induced by a norm.

2.36 Show that in each of the two statements in Theorem 2.3.13, the invariant metric in the conclusion of the statement can be selected to have the nice properties of that conclusion and the further property that the diameter of the space with respect to that metric is at most 1. (Recall that the diameter of a nonempty subset A of a metric space with metric d is $\sup\{ d(x,y) : x,y \in A \}$.)

2.37 (a) Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω and that $0 < p < 1$. Prove that $L_p(\Omega, \Sigma, \mu)$ is locally bounded.

(b) Prove that $L_0[0,1]$ is not locally bounded.

2.38 Prove that no F-space has a countably infinite vector space basis.

2.39 Prove that Zabreĭko's lemma extends to F-spaces, that is, that every countably subadditive seminorm on an F-space is continuous.

2.40 Prove that if a vector space has two topologies under which it is an F-space, and one of the topologies includes the other, then the two topologies are the same.

2.41 (a) Prove the following generalization of Theorem 1.3.14: *Suppose that C is a closed, convex, absorbing subset of an F-space. Then C includes a neighborhood of the origin.* (While this can be done by copying the proof of Theorem 1.3.14 verbatim, there is a shorter proof available. Find it.)

(b) Suppose that X is a vector space with a topology. A *barrel* in X is a closed, convex, balanced, absorbing subset of the space. The space X is *barreled* if each of its barrels includes a neighborhood of the origin. Conclude from (a) that every F-space is barreled.

(c) For some F-spaces, the observation in (b) is not very significant. To see why this is so, list all of the barrels in $L_p[0,1]$ when $0 \le p < 1$.

2.42 Suppose that X is an infinite-dimensional metrizable TVS and that Y is a TVS with more open sets than just $\varnothing$ and Y (which happens, for example, when Y is a Hausdorff TVS such that $Y \neq \{0\}$). Prove that some linear operator from X into Y is unbounded. Conclude that every infinite-dimensional metrizable TVS has a linear functional on it that is unbounded.

2.43 An F-space is often said to have the Hahn-Banach extension property if, for every *closed* subspace of the space, every bounded linear functional on that subspace has an extension as a bounded linear functional to the entire space. Show that this is equivalent to the definition of the Hahn-Banach extension property for F-spaces given near the end of this section, in which the word "closed" is omitted.

2.44 Some texts define an F-space to be a vector space X with a topology induced by a complete invariant metric such that multiplication of vectors by scalars is continuous in each variable separately; that is, which has the property that for each scalar α_0 and each vector x_0, the mappings $f_{\alpha_0}\colon X \to X$ and $g_{x_0}\colon \mathbb{F} \to X$ given by the formulas $f_{\alpha_0}(x) = \alpha_0 x$ and $g_{x_0}(\alpha) = \alpha x_0$ are continuous. The purpose of this exercise is to prove that this definition is equivalent to the one given in Definition 2.3.2. For the moment, let an *F_0-space* be a vector space satisfying the above alternative definition.

(a) Prove that every F-space is an F_0-space.

(b) Suppose that X is an F_0-space whose topology is induced by the complete invariant metric d and that $\epsilon > 0$. Prove that there is a positive δ such that $d(\alpha x, 0) \le \epsilon$ whenever x is a member of X and α a scalar such that $d(x, 0) \le \delta$ and $|\alpha| \le 1$. (Notice that the set $\{ x : x \in X,\ d(\alpha x, 0) \le \epsilon \text{ whenever } |\alpha| \le 1 \}$ has nonempty interior.)

(c) Conclude that every F_0-space is an F-space.

2.4 Topologies Induced by Families of Functions

Suppose that X is a set and that $\mathfrak{F}$ is a family of functions such that each f in $\mathfrak{F}$ maps X into a topological space $(Y_f, \mathfrak{T}_f)$. It is always possible to find a topology for X that makes every member of $\mathfrak{F}$ continuous; for example, just declare every subset of X to be open. However, such a topology might have too many open sets to be of much use, especially when it is important that as many subsets of X as possible be compact. For that reason, it is often desirable to be able to find a topology for X that has just enough open sets to make every member of $\mathfrak{F}$ continuous. As the following result shows, this is always possible.

2.4.1 Proposition. *Let X be a set and let $\mathfrak{F}$ be a family of functions and $\{ (Y_f, \mathfrak{T}_f) : f \in \mathfrak{F} \}$ a family of topological spaces such that each f in $\mathfrak{F}$ maps X into the corresponding Y_f. Then there is a smallest topology for X with respect to which each member of $\mathfrak{F}$ is continuous. That is, there is a unique topology $\mathfrak{T}_{\mathfrak{F}}$ for X such that*

(1) *each f in $\mathfrak{F}$ is $\mathfrak{T}_{\mathfrak{F}}$-continuous; and*

(2) *if $\mathfrak{T}$ is any topology for X such that each f in $\mathfrak{F}$ is $\mathfrak{T}$-continuous, then $\mathfrak{T}_{\mathfrak{F}} \subseteq \mathfrak{T}$.*

The topology $\mathfrak{T}_{\mathfrak{F}}$ has $\{ f^{-1}(U) : f \in \mathfrak{F},\ U \in \mathfrak{T}_f \}$ as a subbasis.

PROOF. Let $\mathfrak{S} = \{ f^{-1}(U) : f \in \mathfrak{F},\ U \in \mathfrak{T}_f \}$ and let $\mathfrak{T}_{\mathfrak{F}}$ be the topology generated by the subbasis $\mathfrak{S}$. Since $\mathfrak{S} \subseteq \mathfrak{T}_{\mathfrak{F}}$, every member of $\mathfrak{F}$ is $\mathfrak{T}_{\mathfrak{F}}$-continuous. Now suppose that $\mathfrak{T}$ is a topology for X such that every member of $\mathfrak{F}$ is $\mathfrak{T}$-continuous. Then $\mathfrak{S} \subseteq \mathfrak{T}$, so $\mathfrak{T}_{\mathfrak{F}} \subseteq \mathfrak{T}$. The uniqueness assertion follows immediately. ∎

2.4.2 Definition. Let all notation be as in the preceding proposition. Then the set $\mathfrak{F}$ is a *topologizing family of functions for X*, and the topology $\mathfrak{T}_{\mathfrak{F}}$ is the *$\mathfrak{F}$ topology of X* or the *topology $\sigma(X, \mathfrak{F})$* or the *weak topology of X induced by $\mathfrak{F}$*. The collection $\{ f^{-1}(U) : f \in \mathfrak{F},\ U \in \mathfrak{T}_f \}$ is the *standard subbasis* for this topology, and the *standard basis* for the topology is the collection of all sets that are intersections of finitely many members of this subbasis.

The term *weak topology* is included in the preceding definition for reference, since it is often given this meaning. However, there is a specific topology for normed spaces called *the* weak topology whose study will be taken up in the next section. To avoid confusion, the term *weak topology* will not be used here in the more general sense.

2.4.3 Example. Let $\{ (X_\alpha, \mathfrak{T}_\alpha) : \alpha \in I \}$ be a family of topological spaces and let X be the Cartesian product $\prod_{\alpha \in I} X_\alpha$. For each α in I, let π_α be the *projection* from X to X_α, that is, the map $x \mapsto x_\alpha$. Let $\mathfrak{F} = \{ \pi_\alpha : \alpha \in I \}$. If $\alpha_0 \in I$ and U is an open subset of X_{α_0}, then $\pi_{\alpha_0}^{-1}(U) = \prod_{\alpha \in I} U_\alpha$, where $U_{\alpha_0} = U$ and $U_\alpha = X_\alpha$ when $\alpha \neq \alpha_0$. Let $\mathfrak{S}_\pi = \{ \pi_\alpha^{-1}(U) : \alpha \in I, U \in \mathfrak{T}_\alpha \}$ and let $\mathfrak{S}_\Pi$ be the standard subbasis for the product topology of X as given in Definition 2.1.5. If $I \neq \varnothing$, then clearly $\mathfrak{S}_\pi = \mathfrak{S}_\Pi$. If $I = \varnothing$, then $\mathfrak{S}_\pi = \varnothing$ but $\mathfrak{S}_\Pi \neq \varnothing$ by the discussion following Definition 2.1.5; however, the two different subbases $\mathfrak{S}_\pi$ and $\mathfrak{S}_\Pi$ both generate the only topology that the one-element set X can have. It follows that the product topology of X is the $\mathfrak{F}$ topology of X, whether or not $I = \varnothing$. That is, the product topology of X is the smallest topology for X with respect to which each of the maps π_α is continuous.

Suppose that X is a topological product and that $\mathfrak{F}$ is the family of projection maps as in the preceding example. Let x be a member of X and (x_β) a net in X; notice that β represents a net index element here, not a member of the index set for the Cartesian product. Then Proposition 2.1.16 says that $x_\beta \to x$ if and only if $f(x_\beta) \to f(x)$ for each f in $\mathfrak{F}$. This result generalizes to arbitrary $\mathfrak{F}$ topologies.

2.4.4 Proposition. *Let X be a set and $\mathfrak{F}$ a topologizing family of functions for X. Suppose that (x_α) is a net in X and x is a member of X. Then $x_\alpha \to x$ with respect to the $\mathfrak{F}$ topology if and only if $f(x_\alpha) \to f(x)$ for each f in $\mathfrak{F}$.*

PROOF. The forward implication follows immediately from the continuity of each member of $\mathfrak{F}$ with respect to the $\mathfrak{F}$ topology. For the converse, suppose that $f(x_\alpha) \to f(x)$ for each f in $\mathfrak{F}$. If $f \in \mathfrak{F}$ and U is a neighborhood of $f(x)$, then there is an $\alpha_{f,U}$ such that $x_\alpha \in f^{-1}(U)$ when $\alpha_{f,U} \preceq \alpha$. It follows from Proposition 2.1.15 that $x_\alpha \to x$ with respect to the $\mathfrak{F}$ topology. ∎

2.4.5 Corollary. *Let W be a topological space, let X be a set topologized by a family $\mathfrak{F}$ of functions, and let g be a function from W into X. Then g is continuous if and only if $f \circ g$ is continuous for each f in $\mathfrak{F}$.*

PROOF. For each f in $\mathfrak{F}$, the continuity of f implies that of $f \circ g$ when g is continuous. Conversely, suppose that $f \circ g$ is continuous for each f in $\mathfrak{F}$. If (w_α) is a net in W that converges to some w, then $f\big(g(w_\alpha)\big) \to f\big(g(w)\big)$ whenever $f \in \mathfrak{F}$, so $g(w_\alpha) \to g(w)$. It follows that g is continuous. ∎

Proposition 2.4.4 is not only a generalization of a statement about net convergence in a topological product, but is in fact itself just a statement about net convergence in a specific topological product. Suppose that X and $\mathfrak{F}$ are as in the statement of the proposition. For each f in $\mathfrak{F}$, let Y_f be the topological space into which f maps X. Because of the equivalence of convergence and coordinatewise convergence for nets in topological products, Proposition 2.4.4 says that a net (x_α) in X converges to some x_0 in X with respect to the $\mathfrak{F}$ topology if and only if $\left(f(x_\alpha)\right)_{f\in\mathfrak{F}} \to \left(f(x_0)\right)_{f\in\mathfrak{F}}$ in the topological product $\prod_{f\in\mathfrak{F}} Y_f$; notice that if $\mathfrak{F}$ is empty, then the notation $\left(f(x)\right)_{f\in\mathfrak{F}}$ is being used to represent the unique element of $\prod_{f\in\mathfrak{F}} Y_f$. It follows immediately that the map $x \mapsto \left(f(x)\right)_{f\in\mathfrak{F}}$ is a homeomorphism from X onto a topological subspace of $\prod_{f\in\mathfrak{F}} Y_f$ *provided the map is one-to-one.* A moment's thought shows that the property needed to assure that the map is one-to-one is exactly the following one.

2.4.6 Definition. Let X be a set and let $\mathfrak{F}$ be a family of functions each of which has domain X. Then the family $\mathfrak{F}$ is *separating* or *total* if, for each pair x and y of distinct elements of X, there is an $f_{x,y}$ in $\mathfrak{F}$ such that $f_{x,y}(x) \neq f_{x,y}(y)$.

The discussion preceding this definition can be summarized by the following result.

2.4.7 Proposition. *Let X be a set and $\mathfrak{F}$ a separating topologizing family of functions for X. For each f in $\mathfrak{F}$, let Y_f be the topological space into which f maps X. Then the map $x \mapsto \left(f(x)\right)_{f\in\mathfrak{F}}$ is a homeomorphism from X with the $\mathfrak{F}$ topology onto a topological subspace of $\prod_{f\in\mathfrak{F}} Y_f$ with the product topology.*

If a topologizing family $\mathfrak{F}$ of functions for a set X is to induce a topology satisfying one of the separation axioms, then it is helpful if each topological space Y_f into which the corresponding member f of $\mathfrak{F}$ maps X satisfies that separation axiom, as will be seen in the next proposition. However, it is far from essential that this be so. For example, let X be any set and let Y be a topological space satisfying none of the separation axioms but having a nonempty proper open subset U. For each x in X, let f_x map x into U and the rest of X into $Y \setminus U$, and let $\mathfrak{F} = \{ f_x : x \in X \}$. Then the $\mathfrak{F}$ topology of X satisfies just about any separation axiom one might devise, since each of its one-element sets, and hence each of its subsets, is open.

On the other hand, the spaces Y_f can satisfy any of the separation axioms whatever, yet the $\mathfrak{F}$ topology will still have no hope of being even a T_0 space if $\mathfrak{F}$ is not a separating family. To see this, suppose that there are distinct elements x_1 and x_2 of X such that $f(x_1) = f(x_2)$ whenever $f \in \mathfrak{F}$. Then a member of the standard subbasis for the $\mathfrak{F}$ topology that contains either x_1

or x_2 must contain both, so the same holds for each member of the standard basis for the $\mathfrak{F}$ topology, from which it follows that the $\mathfrak{F}$ topology is not even T_0.

2.4.8 Proposition. *Let X be a set and $\mathfrak{F}$ a separating topologizing family of functions for X. For each f in $\mathfrak{F}$, let Y_f be the topological space into which f maps X. If each Y_f is T_0, or T_1, or T_2, or T_3, or $T_{3\frac{1}{2}}$, then the $\mathfrak{F}$ topology of X satisfies that same separation axiom.*

PROOF. Throughout this proof, the topology of X is its $\mathfrak{F}$ topology. Suppose that x_1 and x_2 are different elements of X, that f is a member of $\mathfrak{F}$ such that $f(x_1) \neq f(x_2)$, and that U_1 and U_2 are disjoint neighborhoods of $f(x_1)$ and $f(x_2)$ respectively. Then $f^{-1}(U_1)$ and $f^{-1}(U_2)$ are disjoint neighborhoods of x_1 and x_2 respectively. It follows that X is Hausdorff whenever each Y_f is so. The proofs for the T_0 and T_1 axioms are analogous. For the other two cases, recall that topological subspaces and topological products of regular topological spaces are always regular, and that the corresponding result for completely regular spaces is also true; see, for example, [172]. Applications of Proposition 2.4.7 and the fact that homeomorphisms preserve regularity and complete regularity then finish the proof. ■

The preceding result cannot be extended to normal topological spaces, for topological products of normal spaces need not be normal; see [172] for an example. However, the extension to metrizable topological spaces does hold, provided that the topologizing family is countable.

2.4.9 Proposition. *Let X be a set and $\mathfrak{F}$ a separating topologizing family of functions for X. For each f in $\mathfrak{F}$, let Y_f be the topological space into which f maps X. If $\mathfrak{F}$ is countable and the topology of each Y_f is metrizable, then the $\mathfrak{F}$ topology of X is metrizable.*

PROOF. It may be assumed that X has more than one element, and therefore that $\mathfrak{F} \neq \varnothing$. Then the members of $\mathfrak{F}$ can be listed in a sequence (f_n), where the sequence is constant from some term onward if $\mathfrak{F}$ is finite, and the members of $\{ Y_f : f \in \mathfrak{F} \}$ can be listed in a corresponding sequence (Y_n). For each n, let d_n be a metric inducing the topology of Y_n. Define $d \colon X \times X \to [0,1]$ by the formula

$$d(x_1, x_2) = \sum_n \frac{\min\{1, d_n(f_n(x_1), f_n(x_2))\}}{2^n}.$$

It is easy to check that d is a metric on X, and that a net (x_α) in X converges to some x with respect to this metric if and only if $f_n(x_\alpha) \to f_n(x)$ for each n, which happens if and only if $x_\alpha \to x$ with respect to the $\mathfrak{F}$ topology. It follows from Corollary 2.1.22 that the topologies induced on X by d and $\mathfrak{F}$ are the same. ■

2.4.10 Corollary. *If X is a compact topological space and there is a countable separating family of continuous metric-space-valued functions on X, then the topology of X is metrizable.*

PROOF. Let $\mathfrak{F}$ be the countable separating family, let $\mathfrak{T}_{\mathfrak{F}}$ be the $\mathfrak{F}$ topology of X, which is metrizable by the proposition, and let $\mathfrak{T}_c$ be the given compact topology of X. Then $\mathfrak{T}_{\mathfrak{F}} \subseteq \mathfrak{T}_c$ by Proposition 2.4.1. It follows that $\mathfrak{T}_{\mathfrak{F}} = \mathfrak{T}_c$ since Hausdorff topologies are never proper subtopologies of compact topologies; see [200, p. 61] or Exercise 2.8. ■

The countability hypothesis cannot be omitted from Proposition 2.4.9. See Exercise 2.45.

In the rest of this book, almost all of the interest in topologies induced by families of functions will be in topologies induced on a vector space X by subspaces of the vector space $X^\#$ of all linear functionals on X. The following result makes the connection between such topologies and those studied in Section 2.2. The characterization of the dual space in this theorem is a 1940 result of R. S. Phillips [188].

2.4.11 Theorem. *Suppose that X is a vector space and that X' is a subspace of the vector space $X^\#$ of all linear functionals on X. Then the X' topology of X is a locally convex topology, and the dual space of X with respect to this topology is X'.*

PROOF. Throughout this proof, all allusions to a topology for X refer to the X' topology. Suppose that (x_β) and (y_β) are nets in X and (α_β) a net in $\mathbb{F}$ such that all three nets have the same index set, and that (x_β), (y_β), and (α_β) converge to x, y, and α respectively. The continuity of addition and multiplication in $\mathbb{F}$ assures that for each f in X',

$$f(\alpha_\beta x_\beta + y_\beta) = \alpha_\beta f(x_\beta) + f(y_\beta) \to \alpha f(x) + f(y) = f(\alpha x + y),$$

so $\alpha_\beta x_\beta + y_\beta \to \alpha x + y$. It follows that the vector space operations of X are continuous. It is easy to see that

$$\{\, f^{-1}(U) : f \in X', U \text{ is an open ball in } \mathbb{F} \,\}$$

is a subbasis for the topology of X that generates a basis for that topology consisting of convex sets, so X is an LCS.

Let f_0 be a continuous linear functional on X. There is some neighborhood of 0 in X that is mapped by f_0 into the open unit ball of $\mathbb{F}$, from which it follows that there is a nonempty finite collection $f_1, \dots, f_n$ of members of X' and a corresponding collection $U_1, \dots, U_n$ of neighborhoods of 0 in $\mathbb{F}$ such that $f_1^{-1}(U_1) \cap \cdots \cap f_n^{-1}(U_n)$ is mapped by f_0 into the open unit ball of $\mathbb{F}$. Suppose that $x \in \ker(f_1) \cap \cdots \cap \ker(f_n)$. Then $mx \in f_1^{-1}(U_1) \cap \cdots \cap f_n^{-1}(U_n)$ for each positive integer m, and so $m|f_0(x)| = |f_0(mx)| < 1$ whenever $m \in \mathbb{N}$, which implies that $x \in \ker f_0$.

It follows from Lemma 1.9.11 that f_0 is a linear combination of $f_1, \ldots, f_n$, so $f_0 \in X'$. Thus, the dual space of X is included in X'. The reverse inclusion follows from the definition of the X' topology of X, which finishes the proof. ■

It turns out that the topology induced on a vector space X by a subspace of $X^{\#}$ has a particularly simple subbasis and basis.

2.4.12 Proposition. *Suppose that X is a vector space and that X' is a subspace of $X^{\#}$. For each x in X and each f in X', let*

$$B(x, \{f\}) = \{\, y : y \in X,\ |f(y - x)| < 1 \,\}.$$

Similarly, for each x in X and each finite subset A of X', let

$$B(x, A) = \{\, y : y \in X,\ |f(y - x)| < 1 \text{ for each } f \text{ in } A \,\}.$$

Let

$$\mathfrak{S} = \{\, B(x, \{f\}) : x \in X,\ f \in X' \,\}$$

and let

$$\mathfrak{B} = \{\, B(x, A) : x \in X,\ A \text{ is a finite subset of } X' \,\}.$$

Then $\mathfrak{S}$ is a subbasis and $\mathfrak{B}$ a basis for the X' topology of X. If U is a subset of X that is open with respect to the X' topology and x_0 is an element of U, then there is a finite subset A_0 of X' such that $B(x_0, A_0) \subseteq U$; that is, the set U includes a basic neighborhood of x_0 that is "centered" at x_0.

PROOF. Throughout this proof, the topology of X is the X' topology. First of all, notice that for each x in X, each f in X', and each finite subset A of X', the sets $B(x, \{f\})$ and $B(x, A)$ are open. Now let f be a member of X' and U an open subset of $\mathbb{F}$, so that $f^{-1}(U)$ is a member of the standard subbasis for the topology of X, and let x be an element of $f^{-1}(U)$. To show that $\mathfrak{S}$ is a subbasis for the topology of X, it is enough to find a g in X' such that $B(x, \{g\}) \subseteq f^{-1}(U)$. Let ϵ be a positive number such that $\{\, \alpha : \alpha \in \mathbb{F},\ |f(x) - \alpha| < \epsilon \,\} \subseteq U$, and let $g = \epsilon^{-1} f$. Then this g does what is required of it.

Suppose that $f_1, \ldots, f_n \in X'$, that $U_1, \ldots, U_n$ are open subsets of $\mathbb{F}$, and that $x \in f_1^{-1}(U_1) \cap \cdots \cap f_n^{-1}(U_n)$. Then a straightforward modification of the argument just given yields a positive ϵ such that

$$B(x, \{\epsilon^{-1} f_1, \ldots, \epsilon^{-1} f_n\}) \subseteq f_1^{-1}(U_1) \cap \cdots \cap f_n^{-1}(U_n),$$

from which it easily follows that $\mathfrak{B}$ is a basis for the topology of X.

Finally, let x_0 be an element of an open subset U of X, and let x be a member of X and A a finite subset of X' such that $x_0 \in B(x, A) \subseteq U$. It may be assumed that $A \neq \varnothing$. Let δ be such that

$$0 < \delta < 1 - \max\{\,|f(x_0 - x)| : f \in A\,\},$$

and let $A_0 = \delta^{-1}A$. If $y \in B(x_0, A_0)$ and $f \in A$, then

$$|f(y - x)| \leq |f(y - x_0)| + |f(x_0 - x)| < \delta + (1 - \delta) = 1,$$

and so $y \in B(x, A)$. It follows that $x_0 \in B(x_0, A_0) \subseteq B(x, A) \subseteq U$, as required. ■

Let X and X' be as in the preceding two results, with X treated as a topological space under its X' topology. The rest of this section concerns properties that are defined for objects in X because X is a TVS. In particular, the Cauchy condition for nets is defined, and has a characterization analogous to the characterization of net convergence given in Proposition 2.4.4.

2.4.13 Proposition. *Suppose that X is a vector space and that X' is a subspace of $X^\#$. Let (x_α) be a net in X. Then the following are equivalent.*

(a) *The net (x_α) is Cauchy with respect to the X' topology of X.*
(b) *For each f in X', the net $\big(f(x_\alpha)\big)$ is Cauchy in $\mathbb{F}$.*
(c) *For each f in X', the net $\big(f(x_\alpha)\big)$ is convergent in $\mathbb{F}$.*

PROOF. Throughout this proof, the topology of X is the X' topology. It suffices to prove that (a) and (b) are equivalent. Let I be the index set for (x_α). Then $\{\, x_\beta - x_\gamma : (\beta, \gamma) \in I \times I \,\}$ is a net if $I \times I$ is preordered by declaring that $(\beta_1, \gamma_1) \preceq (\beta_2, \gamma_2)$ when $\beta_1 \preceq \beta_2$ and $\gamma_1 \preceq \gamma_2$. Furthermore, it is easy to see that the net (x_α) is Cauchy if and only if the net $(x_\beta - x_\gamma)$ converges to 0; see the discussion preceding Definition 2.1.41. It therefore follows from Proposition 2.4.4 that (x_α) is Cauchy if and only if $\lim_{(\beta,\gamma)}\big(f(x_\beta) - f(x_\gamma)\big) = 0$ for each f in X', which happens if and only if $\big(f(x_\alpha)\big)$ is a Cauchy net in $\mathbb{F}$ for each f in X'. ■

The final two results of this section concern boundedness with respect to the topology induced on a vector space X by a subspace of $X^\#$. The first gives a useful test for such boundedness, while the second shows one dramatic difference between topologies of this sort and norm topologies when the topologizing subspace of $X^\#$ is infinite-dimensional.

2.4.14 Proposition. *Suppose that X is a vector space and that X' is a subspace of $X^\#$. Then a subset A of X is bounded with respect to the X' topology if and only if $f(A)$ is bounded in $\mathbb{F}$ for each f in X'.*

PROOF. Throughout this proof, the topology of X is the X' topology. Let A be a subset of X. Suppose first that A is bounded. Let f be a member of X' and let U be the open unit ball of $\mathbb{F}$. Then there is a positive t such that $A \subseteq tf^{-1}(U)$, which implies that $f(A) \subseteq tU$ and therefore that $f(A)$ is bounded. Conversely, suppose that $f(A)$ is bounded whenever $f \in X'$. Let U_0 be a neighborhood of 0 in X and let $f_1, \ldots, f_n$ be members of X' and $V_1, \ldots, V_n$ neighborhoods of 0 in $\mathbb{F}$ such that $f_1^{-1}(V_1) \cap \cdots \cap f_n^{-1}(V_n) \subseteq U_0$. The boundedness of each $f_j(A)$ yields a positive s such that $f_j(A) \subseteq tV_j$ for each j when $t > s$. It follows that $A \subseteq t\big(f_1^{-1}(V_1) \cap \cdots \cap f_n^{-1}(V_n)\big) \subseteq tU_0$ when $t > s$, so A is bounded. ■

2.4.15 Proposition. *Suppose that X is a vector space and that X' is a subspace of $X^{\#}$. If X' is infinite-dimensional, then every nonempty subset of X that is open with respect to the X' topology is unbounded with respect to that topology.*

PROOF. Throughout this proof, the topology of X is the X' topology. Let U be a nonempty open subset of X. Since translates of bounded subsets of a TVS are bounded, it may be assumed that U is a neighborhood of the origin. Let $f_1, \ldots, f_n$ be members of X' and $V_1, \ldots, V_n$ neighborhoods of 0 in $\mathbb{F}$ such that $f_1^{-1}(V_1) \cap \cdots \cap f_n^{-1}(V_n) \subseteq U$, and let $S = \ker(f_1) \cap \cdots \cap \ker(f_n)$, a subspace of X. It follows that $S \subseteq U$, so it is enough to show that S is unbounded. Let f be a member of X' that is not a linear combination of $f_1, \ldots, f_n$. It follows from Lemma 1.9.11 that $S \setminus \ker(f)$ contains some x. Since $nx \in S$ for each positive integer n and $|f(nx)| = n|f(x)| \to \infty$ as $n \to \infty$, the set $f(S)$ is unbounded. By Proposition 2.4.14, the set S is unbounded. ■

Exercises

2.45 Give an example of a set X and an uncountable separating topologizing family $\mathfrak{F}$ of functions for X such that the $\mathfrak{F}$ topology of X is not metrizable even though each of the topological spaces into which the members of $\mathfrak{F}$ map X is a metric space. Exercise 2.14 might be helpful.

2.46 If X is an infinite-dimensional normed space, then some linear functional on X is not continuous; see Theorem 1.4.11. Show that this result does not generalize to infinite-dimensional Hausdorff locally convex spaces.

2.47 Prove that if a vector space with a topology has a linearly independent sequence in it that converges to 0, then some linear functional on the space is not continuous. Use this fact and your example from the preceding exercise to give an example of an infinite-dimensional Hausdorff LCS in which no linearly independent sequence converges to 0.

2.48 Let X be a set and $\mathfrak{F}$ a separating topologizing family of functions for X. Suppose that for each f in $\mathfrak{F}$, the topological space Y_f into which f maps X is Hausdorff. Prove that X is compact with respect to its $\mathfrak{F}$ topology if and only if $f(X)$ is compact in Y_f for each f in $\mathfrak{F}$ and the image of X in $\prod_{f\in\mathfrak{F}} Y_f$ under the map $x \mapsto (f(x))_{f\in\mathfrak{F}}$ is closed.

2.49 Let X be a vector space with the topology induced by some subspace X' of $X^\#$, and let A be a subset of X. Prove that A is a closed subspace of X if and only if the following holds: For each x in $X \setminus A$, there is an f_x in X' such that $f_x(x) = 1$ and $A \subseteq \ker(f_x)$.

2.50 Let K be a compact Hausdorff space. Show that if $C(K)$ is separable, then the topology of K is metrizable.

2.51 The goal of this exercise is to prove that the vector topologies of a finite-dimensional vector space are exactly the topologies induced by subspaces of the vector space of all linear functionals on that space. This will be done by continuing the argument begun in Exercise 2.34. Let all notation be as in that exercise. All unqualified references to the topologies of X and Y will be to the given topology of X and to the relative topology that Y inherits from that given topology of X.

(a) Prove that the topology of Y is the same as the Y^* topology of Y.

(b) For each y^* in Y^*, each s in S, and each y in Y, let $x^*_{y^*}(s+y) = y^*y$. Show that the map $y^* \mapsto x^*_{y^*}$ is a vector space isomorphism from Y^* onto X^*.

(c) Prove that every open subset of X is open with respect to the X^* topology of X. Conclude that the topology of X is the same as the X^* topology of X.

(d) For each subspace X' of $X^\#$, let $\mathfrak{T}_{X'}$ be the X' topology of X. Show that the map $X' \mapsto \mathfrak{T}_{X'}$ is a one-to-one correspondence between the subspaces of $X^\#$ and the vector topologies of X.

2.52 The preceding exercise says that every vector topology $\mathfrak{T}$ on a finite-dimensional vector space X (which, by Exercise 2.34, is the same as saying every locally convex topology $\mathfrak{T}$ on X) is exactly the topology induced by the vector space X^* of all $\mathfrak{T}$-continuous linear functionals on X. This result clearly does not extend to infinite-dimensional TVSs, since Example 2.2.24 shows that there are Hausdorff TVSs with nonzero elements, and hence other open sets besides the empty set and the entire space, whose dual spaces contain only the zero functional. However, the TVSs in Example 2.2.24 are not LCSs, so this leaves open the possibility that the finite-dimensional result might extend to infinite-dimensional LCSs. The purpose of this exercise is to show that it does not. Prove that if X is an infinite-dimensional normed space, then the norm topology of X is not induced by any subspace of $X^\#$.

2.5 The Weak Topology

The following convention will help prevent a bit of confusion that could otherwise occur throughout the rest of this book. *If X is a normed space, then the notation X^* and the term "the dual space of X" always refer to* *the dual space of X with respect to the norm topology of X, except where explicitly stated otherwise, even in contexts in which another topology for X is being discussed.* Recall also the convention from Section 1.2 concerning topological terms when a normed space is involved: *Whenever reference is made to some topological property in a normed space without specifying the topology, the norm topology is implied.* For example, if a subset of a normed space is said without further qualification to be open, then it is meant that it is open with respect to the norm topology of the space.

It is time to introduce the first of the two topologies that are the main topics of this chapter.

2.5.1 Definition. Let X be a normed space. Then the topology for X induced by the topologizing family X^* is the *weak topology of* X or the X^* *topology of* X or the *topology* $\sigma(X, X^*)$.

That is, the weak topology of a normed space is the smallest topology for the space such that every member of the dual space is continuous with respect to that topology. See Proposition 2.4.1 and Definition 2.4.2.

Since the dual space of a normed space X is a separating family of functions for X by Corollary 1.9.9, and each x^* in X^* maps X into the completely regular space $\mathbb{F}$, it follows from Proposition 2.4.8 that the weak topology of X is itself completely regular. By Theorem 2.4.11, this topology is also locally convex, and the dual space of X with respect to this topology is exactly X^*. As the smallest topology for X with respect to which every member of X^* is continuous, the weak topology of X must be included in every topology with respect to which the members of X^* are all continuous, and in particular must be included in the norm topology of X. These observations are summarized in the following two results.

2.5.2 Theorem. *The weak topology of a normed space is a completely regular locally convex subtopology of the norm topology.*

2.5.3 Proposition. *A linear functional on a normed space is continuous with respect to the weak topology if and only if it is continuous with respect to the norm topology.*

2.5.4 Example. Let (e_n) be the sequence of unit vectors in ℓ_2. It is clear from the usual identification of ℓ_2^* with ℓ_2 that $x^*e_n \to 0$ for each x^* in ℓ_2^*, and so, by Proposition 2.4.4, the sequence (e_n) converges to 0 with respect to the weak topology. Since $\|e_n\|_2 = 1$ for each n, the sequence (e_n) cannot converge to 0 with respect to the norm topology. The norm and weak topologies of ℓ_2 are therefore different, so it is possible for the weak topology of a normed space to be a proper subtopology of the norm topology.

A topological property that holds with respect to the weak topology is said to be a *weak* property or to hold *weakly*. For example, the preceding theorem and proposition could be restated in part by saying that weakly open subsets of a normed space are always open and that continuity and weak continuity are equivalent for linear functionals on a normed space. Attaching the letter w to a topological symbol is another way to indicate that the weak topology is being used. For example, the weak convergence of a net (x_α) to an element x can be expressed by writing $x_\alpha \xrightarrow{w} x$ or $w\text{-}\lim_\alpha x_\alpha = x$, while the weak closure of a set A can be denoted by $\overline{A}^w$.

Of course, all of the results derived in Sections 2.2 and 2.4 for a Hausdorff locally convex topology induced by a separating vector space of linear functionals hold for the weak topology of a normed space X. In particular, it follows from Propositions 2.4.4 and 2.4.13 that if (x_α) is a net in X and x is an element of X, then $x_\alpha \xrightarrow{w} x$ if and only if $x^*x_\alpha \to x^*x$ for each x^* in X^*, and (x_α) is weakly Cauchy if and only if (x^*x_α) is a Cauchy (that is, convergent) net in $\mathbb{F}$ for each x^* in X^*. A particularly useful basis for the weak topology of X is given by the collection of all sets of the form

$$\{\, y : y \in X,\ |x^*(y - x)| < 1 \text{ for each } x^* \text{ in } A \,\}$$

such that $x \in X$ and A is a finite subset of X^*; see Proposition 2.4.12 for the development of this basis and some other information about it. The reader might also wish to review the properties of a TVS given in Theorem 2.2.9, since those properties will be used extensively in what is to follow.

As a special case of Definition 2.2.8, a subset A of a normed space X is *weakly bounded* if, for each weak neighborhood U of 0 in X, there is a positive s_U such that $A \subseteq tU$ whenever $t > s_U$. By Proposition 2.4.14, this is equivalent to requiring that $x^*(A)$ be a bounded set of scalars for each x^* in X^*. As Example 2.5.4 shows, the weak topology of a normed space can be a proper subtopology of the norm topology, so it might seem easier for a subset of a normed space to be weakly bounded than to be bounded. Perhaps surprisingly, this is not the case.

2.5.5 Theorem. *A subset of a normed space is bounded if and only if it is weakly bounded.*

PROOF. It is clear that every bounded subset of a normed space is weakly bounded, since every weakly open subset of a normed space is open. Conversely, suppose that A is a weakly bounded subset of a normed space X. It may be assumed that A is nonempty. Let Q be the natural map from X into X^{**}. Then $Q(A)$ is a nonempty collection of bounded linear functionals on the Banach space X^*. For each x^* in X^*,

$$\sup\{\, |(Qx)(x^*)| : x \in A \,\} = \sup\{\, |x^*x| : x \in A \,\} < \infty.$$

It follows from the uniform boundedness principle that

$$\sup\{\, \|x\| : x \in A \,\} = \sup\{\, \|Qx\| : x \in A \,\} < \infty,$$

as required. ∎

2.5.6 Corollary. *A subset A of a normed space X is bounded if and only if $x^*(A)$ is a bounded set of scalars for each x^* in X^*.*

As special cases of Propositions 2.2.11 and 2.2.12, every weakly compact set and every weakly Cauchy sequence in a normed space is weakly bounded, and every weakly convergent sequence in a normed space is weakly Cauchy by Proposition 2.1.47. The next two corollaries of the above theorem follow immediately.

2.5.7 Corollary. *Weakly compact subsets of a normed space are bounded.*

2.5.8 Corollary. *In a normed space, weakly Cauchy sequences, and so weakly convergent sequences, are bounded.*

2.5.9 Corollary. *Every nonempty weakly open subset of an infinite-dimensional normed space is unbounded.*

PROOF. Let X be an infinite-dimensional normed space. Then X^* is also infinite-dimensional. By Proposition 2.4.15, every nonempty weakly open subset of X is weakly unbounded, and therefore unbounded. ■

Theorem 2.5.5 and its corollaries have several important consequences for the continuity of linear operators. Suppose that T is a linear operator from a normed space X into a normed space Y. Then T is bounded if and only if the set $T(B_X)$ is bounded, which by Corollary 2.5.6 happens if and only if $y^*T(B_X)$ is bounded for each y^* in Y^*. This immediately yields the following result.

2.5.10 Proposition. *A linear operator T from a normed space X into a normed space Y is bounded if and only if $y^*T \in X^*$ whenever $y^* \in Y^*$.*

Since a linear functional on a normed space is continuous if and only if it is weakly continuous, the above proposition also says that the linear operator T is continuous if and only if y^*T is a weakly continuous linear functional on X whenever $y^* \in Y^*$. By Corollary 2.4.5, this is equivalent to the weak-to-weak continuity of T.

2.5.11 Theorem. *A linear operator T from a normed space X into a normed space Y is norm-to-norm continuous if and only if it is weak-to-weak continuous.*

2.5.12 Corollary. *A linear operator T from a normed space X onto a normed space Y is an isomorphism of normed spaces if and only if it is a weak-to-weak homeomorphism.*

In particular, if two normed spaces are implicitly treated as the same because of some natural isometric isomorphism from one onto the other, such as the way c_0^* is usually identified with ℓ_1, then the weak topologies of the two spaces are preserved by the same isometric isomorphism.

If X is an infinite-dimensional normed space, then Corollary 2.5.9 implies that its open unit ball cannot be weakly open, so the norm and weak topologies of X must differ. This could not happen if X were finite-dimensional, for then the norm and weak topologies of X would both be the unique Hausdorff vector topology of X and so would be the same; see Theorem 2.2.31. Summarizing these observations yields the following useful characterization of normed spaces whose norm and weak topologies are identical.

2.5.13 Proposition. *The norm and weak topologies of a normed space are the same if and only if the space is finite-dimensional.*

The weak topology of an infinite-dimensional normed space is therefore not induced by the norm of the space. In fact, the weak topology of such a space is not induced by any metric at all.

2.5.14 Proposition. *The weak topology of a normed space is induced by a metric if and only if the space is finite-dimensional.*

PROOF. The weak topology of a finite-dimensional normed space is the same as the norm topology, and so is induced by a metric. For the converse, suppose that the weak topology of some infinite-dimensional normed space X is induced by a metric d. For each positive integer n, let U_n be the d-open ball in X centered at 0 having d-radius n^{-1}. By Corollary 2.5.9, each U_n contains some x_n such that $\|x_n\| \geq n$. It follows that (x_n) is an unbounded sequence in X that converges weakly to 0, which contradicts Corollary 2.5.8. ■

Completeness is yet another property that the weak topology cannot have unless the normed space is finite-dimensional.

2.5.15 Proposition. *The weak topology of a normed space is complete if and only if the space is finite-dimensional.*

PROOF. Let X be a normed space. If X is finite-dimensional, then X is a Banach space that is the same TVS with respect to the weak topology as it is with respect to the norm topology, so the weak topology of X is complete by Corollary 2.1.51.

Suppose that X is infinite-dimensional. Since X^* is infinite-dimensional, it follows from Theorem 1.4.11 that some linear functional f on X^* is unbounded. It will first be shown that if F_0 is a finite subset of X^*, then there is an x_{F_0} in X such that $x^*x_{F_0} = fx^*$ for each x^* in F_0. This is trivial if F_0 is empty, so it may be assumed that F_0 is a nonempty finite subset $\{x_1^*, \ldots, x_n^*\}$ of X^*. Let M be the norm of the bounded linear functional obtained by restricting f to the finite-dimensional subspace $\langle\{x_1^*, \dots, x_n^*\}\rangle$ of X^*. For each linear combination $\alpha_1 x_1^* + \cdots + \alpha_n x_n^*$ of $x_1^*, \dots, x_n^*$,

$$\begin{aligned}|\alpha_1 f x_1^* + \cdots + \alpha_n f x_n^*| &= |f(\alpha_1 x_1^* + \cdots + \alpha_n x_n^*)| \\ &\leq M \|\alpha_1 x_1^* + \cdots + \alpha_n x_n^*\|.\end{aligned}$$

By Helly's theorem, there is an x_{F_0} in X such that $x_j^* x_{F_0} = f x_j^*$ when $j = 1, \dots, n$, as was claimed.

Let I be the collection of all finite subsets of X^*, preordered by declaring that $F_1 \preceq F_2$ when $F_1 \subseteq F_2$. For each F in I, let x_F be a member of X such that $x^* x_F = f x^*$ for each x^* in F. Then the net $(x_F)_{F \in I}$ is weakly Cauchy, since $x^* x_F = f x^*$ whenever $x^* \in X^*$ and $\{x^*\} \preceq F$. However, the net (x_F) is not weakly convergent, for if it were to converge weakly to some x in X, then it would follow that $f x^* = \lim_F x^* x_F = x^* x$ for each x^* in X^*, which would imply that f is in the image of X in X^{**} under the natural map, in contradiction to the fact that f is unbounded. ■

Suppose that X is an infinite-dimensional normed space. Then X must have open convex subsets that are not weakly open; in particular, the open balls in X cannot be weakly open since they are bounded. Because of this, it might not seem likely that the closed convex subsets of an arbitrary normed space would have to be the same as its weakly closed convex subsets. However, Corollary 2.2.29 assures that this must be the case, since the dual space of a normed space with respect to the norm topology is the same as the dual space with respect to the weak topology. The following theorem is just a special case of that corollary.

2.5.16 Theorem. (S. Mazur, 1933 [162]). *The closure and weak closure of a convex subset of a normed space are the same. In particular, a convex subset of a normed space is closed if and only if it is weakly closed.*

Mazur actually showed that every closed convex subset of a normed space is weakly sequentially closed, but his method can be used to prove the more general result.

2.5.17 Corollary. (S. Banach, 1932 [13]). *The closure and weak closure of a subspace of a normed space are the same, so a subspace of a normed space is closed if and only if it is weakly closed.*

2.5.18 Corollary. *If A is a subset of a normed space, then $\overline{\mathrm{co}}(A) = \overline{\mathrm{co}}^{\,w}(A)$.*

PROOF. It follows from Theorem 2.2.9 (i) that $\overline{\mathrm{co}}(A) = \overline{\mathrm{co}(A)} = \overline{\mathrm{co}(A)}^{\,w} = \overline{\mathrm{co}}^{\,w}(A)$. ■

2.5.19 Corollary. *If $(x_\alpha)_{\alpha\in I}$ is a net in a normed space that converges weakly to some x, then some sequence of convex combinations of members of $\{x_\alpha : \alpha \in I\}$ converges to x with respect to the norm topology.*

PROOF. Since $x \in \overline{\text{co}}(\{x_\alpha : \alpha \in I\})$ by the preceding corollary, some sequence in $\text{co}(\{x_\alpha : \alpha \in I\})$ must converge to x. ■

One of the basic properties of normed spaces is that the norm function $x \mapsto \|x\|$ is continuous. However, it does not have to be weakly continuous. For instance, the norm function on ℓ_2 is not weakly continuous since, by Example 2.5.4, there are sequences in the unit sphere of ℓ_2 that converge weakly to 0. In fact, it can be shown that the norm function is weakly continuous if and only if the norm and weak topologies of the space are the same, that is, if and only if the space is finite-dimensional. See Exercise 2.54.

Thus, it is not always true that $\|x_\alpha\| \to \|x\|$ when a net (x_α) in a normed space converges weakly to some x. Something can be said in this situation, but saying it requires the definition of the limit inferior of a net of real numbers.

Suppose that $(t_\alpha)_{\alpha\in I}$ is a net of real numbers. For each α in I, let $i_\alpha = \inf\{t_\beta : \alpha \preceq \beta\}$ and $s_\alpha = \sup\{t_\beta : \alpha \preceq \beta\}$; notice that i_α and s_α could be infinite. If $\alpha_1 \preceq \alpha_2$, then $i_{\alpha_1} \leq i_{\alpha_2}$ and $s_{\alpha_1} \geq s_{\alpha_2}$, from which it follows that $\lim_\alpha i_\alpha$ and $\lim_\alpha s_\alpha$ both exist, provided that the notion of the limit of a net of real numbers is extended in the obvious way to nets of extended real numbers, that is, to nets in the ordered set $\mathbb{R}\cup\{-\infty, +\infty\}$. By analogy with sequences of real numbers, it is natural to call these limits the limit inferior and limit superior, respectively, of (t_α).

2.5.20 Definition. Let $(t_\alpha)_{\alpha\in I}$ be a net in $\mathbb{R}$. Then $\lim_\alpha \inf\{t_\beta : \alpha \preceq \beta\}$ and $\lim_\alpha \sup\{t_\beta : \alpha \preceq \beta\}$ are, respectively, the *limit inferior* and *limit superior* of (t_α), and are denoted, respectively, by $\liminf_\alpha t_\alpha$ and $\limsup_\alpha t_\alpha$.

See Exercise 2.55 for some properties of the limit inferior and limit superior that can be extended from sequences to nets.

2.5.21 Theorem. *If (x_α) is a weakly convergent net in a normed space, then $\|w\text{-}\lim_\alpha x_\alpha\| \leq \liminf_\alpha \|x_\alpha\|$.*

PROOF. Let x be the weak limit of a weakly convergent net $(x_\alpha)_{\alpha\in I}$ in a normed space X. It may be assumed that $x \neq 0$. By Corollary 1.9.8, there is an x^* in S_{X^*} such that $x^*x = \|x\|$, so $\|x\| = \lim_\alpha |x^*x_\alpha|$. For each positive ϵ, there is an α_ϵ in I such that $\|x\| - \epsilon \leq |x^*x_\alpha| \leq \|x_\alpha\|$ if $\alpha_\epsilon \preceq \alpha$, from which it follows that $\|x\| \leq \liminf_\alpha \|x_\alpha\|$. ■

A function f from a topological space X into $\mathbb{R}$ is said to be *lower semicontinuous* if $f(x) \leq \liminf_\alpha f(x_\alpha)$ whenever (x_α) is a net in X converging to some element x of X. Thus, the preceding theorem says that norm functions are weakly lower semicontinuous.

Suppose that M is a subspace of a normed space X. Then a statement such as "the net (x_α) in M converges weakly to an x in M" might seem ambiguous, since it is not clear whether the statement refers to the weak topology of X or to that of M treated as a normed space in its own right. Fortunately, it makes no difference.

2.5.22 Proposition. *Let M be a subspace of a normed space X. Then the weak topology of the normed space M is the same as the topology of M inherited from the weak topology of X.*

PROOF. Let (x_α) be a net in M and x an element of M. It is enough to show that (x_α) converges to x with respect to the weak topology of M if and only if (x_α) converges to x with respect to the weak topology of X. Since the restriction of a member of X^* to M always lies in M^*, and each member m^* of M^* has a Hahn-Banach extension $x^*_{m^*}$ in X^*, it follows that $x^*x_\alpha \to x^*x$ for each x^* in X^* if and only if $m^*x_\alpha \to m^*x$ for each m^* in M^*, as required. ∎

The rest of this section deals with sequential properties of the weak topology and ways in which those sequential properties can mimic those of the norm topology.

It follows from Proposition 2.5.15 that every infinite-dimensional normed space contains a weakly Cauchy net with no weak limit, but that does not eliminate the possibility that all of the weakly Cauchy *sequences* in such a space could be weakly convergent. This sometimes happens.

2.5.23 Definition. A normed space is *weakly sequentially complete* if every weakly Cauchy sequence in the space has a weak limit.

2.5.24 Example. Let $\big((\alpha_n^{(k)})\big)_{k=1}^\infty$ be a weakly Cauchy sequence in ℓ_1. For each positive integer n_0, the map $(\beta_n) \mapsto \beta_{n_0}$ is a bounded linear functional on ℓ_1, which implies that the sequence $(\alpha_{n_0}^{(k)})_{k=1}^\infty$ of n_0^{th} terms of the members of $\big((\alpha_n^{(k)})\big)_{k=1}^\infty$ has a limit α_{n_0}. It will be shown that the sequence (α_n) of termwise limits of $\big((\alpha_n^{(k)})\big)_{k=1}^\infty$ is itself a member of ℓ_1 and that $\|(\alpha_n^{(k)}) - (\alpha_n)\|_1 \to 0$.

To this end, let M be a bound for the norms of the members of the sequence $\big((\alpha_n^{(k)})\big)_{k=1}^\infty$. Then for each pair of positive integers m and k,

$$\sum_{n=1}^m |\alpha_n^{(k)}| \le M.$$

Treating m as fixed and letting k tend to ∞ in this inequality shows that

$$\sum_{n=1}^m |\alpha_n| \le M$$

for each positive integer m, from which it follows that

$$\sum_{n=1}^{\infty} |\alpha_n| \leq M,$$

that is, that $(\alpha_n) \in \ell_1$.

Suppose that $\|(\alpha_n^{(k)}) - (\alpha_n)\|_1$ does not tend to 0. Then there is a subsequence $\left((\alpha_n^{(k_j)})\right)_{j=1}^{\infty}$ of $\left((\alpha_n^{(k)})\right)_{k=1}^{\infty}$ and a positive scalar t such that

$$\|t(\alpha_n^{(k_j)}) - t(\alpha_n)\|_1 \geq 1$$

for each j. Let $\beta_n^{(j)} = t\alpha_n^{(k_j)} - t\alpha_n$ for each j and each n. Then $\left((\beta_n^{(j)})\right)_{j=1}^{\infty}$ is a weakly Cauchy sequence in ℓ_1 such that $\|(\beta_n^{(j)})\|_1 \geq 1$ for each j and $\lim_j \beta_n^{(j)} = 0$ for each n. After thinning $\left((\beta_n^{(j)})\right)_{j=1}^{\infty}$ if necessary, it may be assumed that there is a sequence (n_j) of nonnegative integers such that $0 = n_1 < n_2 < \cdots$ and

$$\sum_{m=n_j+1}^{n_{j+1}} |\beta_m^{(j)}| > \tfrac{3}{4}\|(\beta_n^{(j)})\|_1$$

for each j. For each positive integer m, let γ_m be a scalar of absolute value 1 such that if j is the positive integer for which $n_j < m \leq n_{j+1}$, then $\gamma_m\beta_m^{(j)} = (-1)^j|\beta_m^{(j)}|$. Let x^* be the member of ℓ_1^* that is represented in the usual way by the member (γ_n) of ℓ_∞. For each positive even integer j,

$$\begin{aligned}
\operatorname{Re} x^*(\beta_n^{(j)}) &= \sum_{m=n_j+1}^{n_{j+1}} \gamma_m\beta_m^{(j)} + \sum_{m\in\mathbb{N}\setminus\{n_j+1,\dots,n_{j+1}\}} \operatorname{Re}(\gamma_m\beta_m^{(j)}) \\
&\geq \sum_{m=n_j+1}^{n_{j+1}} |\beta_m^{(j)}| - \sum_{m\in\mathbb{N}\setminus\{n_j+1,\dots,n_{j+1}\}} |\beta_m^{(j)}| \\
&> \tfrac{3}{4}\|(\beta_n^{(j)})\|_1 - \tfrac{1}{4}\|(\beta_n^{(j)})\|_1 \\
&\geq \tfrac{1}{2},
\end{aligned}$$

and similarly $\operatorname{Re} x^*(\beta_n^{(j)}) < -\frac{1}{2}$ whenever j is odd. This contradicts the weak Cauchyness of $\left((\beta_n^{(j)})\right)_{j=1}^{\infty}$, so it must be that $\|(\alpha_n^{(k)}) - (\alpha_n)\|_1 \to 0$.

Thus, every weakly Cauchy sequence in ℓ_1 is norm convergent, from which it follows that ℓ_1 is weakly sequentially complete even though, by Proposition 2.5.15, it cannot be weakly complete.

Since weakly convergent sequences are weakly Cauchy, it also follows that every weakly convergent sequence in ℓ_1 is actually norm convergent to the weak limit of the sequence. The fact that ℓ_1 possesses this property first appeared in a 1920 paper by J. Schur [212], and is important enough that Schur's name has become attached to the property.

2.5.25 Definition. A normed space has *Schur's property* if it satisfies the following condition: Whenever (x_n) is a sequence in the space and x an element of the space such that $x_n \xrightarrow{w} x$, it follows that $x_n \to x$.

By Example 2.5.4, the space ℓ_2 does not have Schur's property. As will be shown, it does have the following weakened version of the property.

2.5.26 Definition. A normed space has the *Radon-Riesz property* or the *Kadets-Klee property* or *property* (H), and is called a *Radon-Riesz space,* if it satisfies the following condition: Whenever (x_n) is a sequence in the space and x an element of the space such that $x_n \xrightarrow{w} x$ and $\|x_n\| \to \|x\|$, it follows that $x_n \to x$.

2.5.27 Example. Let (x_k) be a sequence in ℓ_2 and x an element of ℓ_2 such that $x_k \xrightarrow{w} x$ and $\|x_k\|_2 \to \|x\|_2$. It will be shown that $x_k \to x$. For each k, let $(\alpha_n^{(k)})$ be the sequence of scalars that is x_k and let $\overline{x_k}$ be the member of ℓ_2 obtained by replacing each scalar α in that sequence with its complex conjugate $\overline{\alpha}$. Similarly, let (α_n) be the sequence of scalars that is x and let $\overline{x} = (\overline{\alpha_n})$. Let x^* and $\overline{x}^*$ be the members of ℓ_2^* that are identified in the usual way with x and $\overline{x}$ respectively. Then

$$\begin{aligned}
\|x_k - x\|_2^2 &= \sum_n (\alpha_n^{(k)} - \alpha_n)\overline{(\alpha_n^{(k)} - \alpha_n)} \\
&= \sum_n |\alpha_n^{(k)}|^2 - \sum_n \alpha_n \overline{\alpha_n^{(k)}} - \sum_n \overline{\alpha_n}\alpha_n^{(k)} + \sum_n |\alpha_n|^2 \\
&= \|x_k\|_2^2 - x^*\overline{x_k} - \overline{x}^* x_k + \|x\|_2^2 \\
&\to \|x\|_2^2 - x^*\overline{x} - \overline{x}^* x + \|x\|_2^2 \\
&= 0.
\end{aligned}$$

Thus, the space ℓ_2 has the Radon-Riesz property.

The reason that the Radon-Riesz property is named after J. Radon and F. Riesz is that they proved that the spaces $L_p(\Omega, \Sigma, \mu)$, where μ is a positive measure on a σ-algebra Σ of subsets of a set Ω, have it when $1 < p < \infty$, with Radon's proof coming in 1913 [192] and Riesz's in 1928–1929 [196, 197]. M. I. Kadets and V. L. Klee used versions of the Radon-Riesz property to develop the ingredients for the proof that all infinite-dimensional separable Banach spaces are homeomorphic; see [119], [138], and [122]. In addition, Kadets used it in [120] to prove that every separable normed space has an equivalent locally uniformly rotund norm; the definition of this rotundity property will be given in Chapter 5. Because of the work that Kadets and Klee did with the Radon-Riesz property, it is sometimes named after them. The Radon-Riesz property is also sometimes called property (H) since a strengthened form of it was given that name and studied in a 1958 paper by K. Fan and I. Glicksberg [77], and M. M. Day later adopted that notation for the Radon-Riesz property in his book [56]. Incidentally, the letter H in the notation does not stand for anything. The Fan and Glicksberg paper has an alphabetically-labeled list of properties for normed spaces that starts with (A) and ends with (H), and their strengthened version of the Radon-Riesz property happens to fall last in the list. Such is often the way mathematical notation gets established!

Exercises

2.53 Suppose that X is c_0 or ℓ_p, where $1 < p < \infty$. Let $\big((\beta_n^{(\alpha)})\big)_{\alpha\in I}$ be a net in X and (β_n) an element of X.

(a) Show that if $(\beta_n^{(\alpha)}) \xrightarrow{w} (\beta_n)$, then $\beta_n^{(\alpha)} \to \beta_n$ for each n.

(b) Show that if the net $\big((\beta_n^{(\alpha)})\big)_{\alpha\in I}$ is bounded and $\beta_n^{(\alpha)} \to \beta_n$ for each n, then $(\beta_n^{(\alpha)}) \xrightarrow{w} (\beta_n)$.

(c) Show that the conclusion of (b) can fail if the net $\big((\beta_n^{(\alpha)})\big)_{\alpha\in I}$ is not required to be bounded.

(d) Suppose that $\big((\beta_n^{(\alpha)})\big)_{\alpha\in I}$ is a net in ℓ_1 and (β_n) an element of ℓ_1. Does the result stated in (a) still hold? What about (b)?

2.54 Suppose that X is an infinite-dimensional normed space. Show that there is a net in S_X that converges weakly to 0. (Notice that this implies that the map $x \mapsto \|x\|$ from a normed space into $\mathbb{F}$ is not weakly continuous if the normed space is infinite-dimensional.)

2.55 For the purposes of this exercise, the notions of infimum, supremum, and net limit in $\mathbb{R}$ have been extended in the obvious way to the extended real numbers. Suppose that (t_α) and (u_α) are nets in $\mathbb{R}$ having the same index set I. Prove that the following hold.

(a) $\liminf_\alpha t_\alpha = \sup\{\inf\{t_\beta : \alpha \preceq \beta\} : \alpha \in I\}$ and $\limsup_\alpha t_\alpha = \inf\{\sup\{t_\beta : \alpha \preceq \beta\} : \alpha \in I\}$.

(b) $\liminf_\alpha t_\alpha \le \limsup_\alpha t_\alpha$.

(c) If $s \ge 0$, then $\liminf_\alpha(st_\alpha) = s\liminf_\alpha t_\alpha$ and $\limsup_\alpha(st_\alpha) = s\limsup_\alpha t_\alpha$.

(d) If $s \le 0$, then $\liminf_\alpha(st_\alpha) = s\limsup_\alpha t_\alpha$ and $\limsup_\alpha(st_\alpha) = s\liminf_\alpha t_\alpha$.

(e) $\liminf_\alpha t_\alpha + \liminf_\alpha u_\alpha \le \liminf_\alpha(t_\alpha + u_\alpha)$ and $\limsup_\alpha(t_\alpha + u_\alpha) \le \limsup_\alpha t_\alpha + \limsup_\alpha u_\alpha$ except in cases in which a sum of lim inf's or lim sup's is formally $+\infty - \infty$ or $-\infty + \infty$ and therefore not defined.

(f) There are subnets (t_γ) and (t_δ) of (t_α) such that $\lim_\gamma t_\gamma = \liminf_\alpha t_\alpha$ and $\lim_\delta t_\delta = \limsup_\alpha t_\alpha$.

(g) If (t_ζ) is a convergent subnet of (t_α), then $\liminf_\alpha t_\alpha \le \lim_\zeta t_\zeta \le \limsup_\alpha t_\alpha$. (Notice that $\lim_\zeta t_\zeta$ could be $\pm\infty$.)

(h) If (t_ζ) is any subnet of (t_α) whatever, then $\liminf_\alpha t_\alpha \le \liminf_\zeta t_\zeta \le \limsup_\zeta t_\zeta \le \limsup_\alpha t_\alpha$.

(i) $\liminf_\alpha t_\alpha = \limsup_\alpha t_\alpha$ if and only if $\lim_\alpha t_\alpha$ exists. If $\lim_\alpha t_\alpha$ exists, then $\lim_\alpha t_\alpha = \liminf_\alpha t_\alpha = \limsup_\alpha t_\alpha$.

2.56 Prove that a normed space with its weak topology is of the second category in itself if and only if the space is finite-dimensional. (The fact that an infinite-dimensional normed space with its weak topology is of the first category in itself is a 1938 result of J. V. Wehausen [241].)

2.57 Find a one-to-one linear operator from a Hausdorff TVS onto a normed space such that the operator is bounded but not continuous. (Compare this to Theorems 1.4.2 and 2.3.28.)

2.58 Prove that if ℓ_1 is isomorphically embedded in a normed space X, then some bounded sequence in X has no weakly Cauchy subsequence. (Compare Exercise 2.66.) This is the easy part of *Rosenthal's ℓ_1 theorem* (H. P. Rosenthal, 1974 [199]): *Each bounded sequence in a Banach space X has a weakly Cauchy subsequence if and only if ℓ_1 is not isomorphically embedded in X.* See Joseph Diestel's book [58] for an extensive discussion of this theorem and its consequences.

2.59 Show that c_0 is not weakly sequentially complete.

2.60 Suppose that K is a compact Hausdorff space. Show that $C(K)$ is weakly sequentially complete if and only if K is finite.

2.61 Suppose that $1 < p < \infty$. Show that ℓ_p is weakly sequentially complete. (Do not use any results from later sections of this book to do this. As will be seen in Section 2.8, every reflexive normed space is weakly sequentially complete, but that result is a corollary of a fairly deep theorem of that section.)

2.62 Show that the spaces c_0 and ℓ_p such that $1 < p < \infty$ all lack Schur's property.

2.63 Show that if a normed space X has Schur's property, then so does every subspace of X, but in general the same cannot be guaranteed of every quotient space X/M such that M is a closed subspace of X.

2.64 Show that a normed space X has the Radon-Riesz property if and only if it satisfies this condition: Whenever (x_n) is a sequence in S_X and x is an element of S_X such that $x_n \xrightarrow{w} x$, it follows that $x_n \to x$.

2.65 A normed space X has the *semi-Radon-Riesz property* if it satisfies the following condition: Whenever (x_n), x, and (x_n^*) are, respectively, a sequence in S_X, an element of S_X, and a sequence in S_{X^*} such that $x_n^* x_n = 1$ for each n and $x_n \xrightarrow{w} x$, it follows that $x_n^* x \to 1$.

(a) Show that the Radon-Riesz property implies the semi-Radon-Riesz property.

(b) Show that c_0 does not have the semi-Radon-Riesz property.

The definition of this property is due to L. P. Vlasov [236], who called it property (SA) and provided an example to show that it is properly weaker than the Radon-Riesz property. See also Vlasov [238] and Megginson [168] for further studies of the property.

2.66 This exercise assumes familiarity with ultranets; see Appendix D. Prove that every bounded ultranet in a normed space is weakly Cauchy. Conclude from this that every bounded net in a normed space has a weakly Cauchy subnet. (Compare Exercise 2.58.)

2.6 The Weak* Topology

The dual space of a normed space is the setting for the second of the two topologies that are the main focus of this chapter.

2.6.1 Definition. Let X be a normed space and let Q be the natural map from X into X^{**}. Then the topology for X^* induced by the topologizing family $Q(X)$ is the *weak** (pronounced "weak star") *topology of* X^* or the *X topology of* X^* or the *topology* $\sigma(X^*, X)$.

That is, the weak* topology of the dual space of a normed space X is the smallest topology for X^* such that, for each x in X, the linear functional $x^* \mapsto x^*x$ on X^* is continuous with respect to that topology.

By analogy with the weak topology, a topological property that holds with respect to the weak* topology is said to hold *weakly**[5] or to be a *weak** property. Whenever w^* is attached to a topological symbol, it indicates that the reference is to the weak* topology. Examples of this would be $x_\alpha^* \xrightarrow{w^*} x^*$, $w^*\text{-}\lim_\alpha x_\alpha^* = x^*$, and $\overline{A}^{w^*}$.

Let X and Q be as in Definition 2.6.1. If x^* and y^* are different elements of X^*, then there is an x in X such that $x^*x \neq y^*x$, so $Q(X)$ is a separating family of functions for X^*. It therefore follows from Proposition 2.4.8 and Theorem 2.4.11 that the weak* topology of X^* is a completely regular locally convex topology and that the dual space of X^* with respect to this topology is $Q(X)$. Since $Q(X) \subseteq X^{**}$ and the weak topology of X^* is the smallest topology for X^* with respect to which all members of X^{**} are continuous, the weak* topology of X^* is included in the weak topology of X^*, and the two topologies are the same if and only if $Q(X) = X^{**}$, that is, if and only if X is reflexive. The following theorem and proposition summarize these observations.

---

[5] "Weakly*" might sound somewhat awkward, but the alternatives are at least equally unpleasant. "Weak*ly" is perhaps grammatically correct, but too outrageous to consider. "Weak*" is sometimes used as both an adjective and adverb, but consistency then demands that "weak" be used as an adverb, and calling a function "weak continuous" also seems awkward. "Weakly*" was suggested to the author by Mahlon Day.

2.6.2 Theorem. *Let X be a normed space. Then the weak* topology of X^* is a completely regular locally convex subtopology of the weak topology of X^*, and therefore of the norm topology of X^*. Furthermore, the weak* and weak topologies of X^* are the same if and only if X is reflexive.*

2.6.3 Corollary. *Let X be a normed space. Then the weak* and norm topologies of X^* are the same if and only if X is finite-dimensional.*

PROOF. If X is infinite-dimensional, then the weak topology of X^* is a proper subtopology of the norm topology by Proposition 2.5.13, and therefore so must be the weak* topology. If X is finite-dimensional, then the weak*, weak, and norm topologies of X^* are each the unique Hausdorff vector topology of X^*. ■

2.6.4 Proposition. *Let X be a normed space. Then a linear functional on X^* is weakly* continuous if and only if it has the form $x^* \mapsto x^*x_0$ for some x_0 in X.*

2.6.5 Example. Let (e_n^*) be the sequence of elements of c_0^* that correspond in the usual way to the standard unit vectors of ℓ_1. It is clear that (e_n^*) is weakly* convergent to 0. By Corollary 2.5.12, the natural isometric isomorphism from ℓ_1 onto c_0^* is also a weak-to-weak homeomorphism, from which it follows that (e_n^*) does not converge weakly to 0. The weak and weak* topologies of c_0^* are therefore different, as Theorem 2.6.2 assures must be the case since c_0 is not reflexive.

As with the weak topology of a normed space, the results obtained in Sections 2.2 and 2.4 for a Hausdorff locally convex topology induced by a separating vector space of linear functionals all hold for the weak* topology of the dual space X^* of a normed space X. For example, Propositions 2.4.4 and 2.4.13 imply that if (x_α^*) is a net in X^* and x^* is a member of X^*, then (x_α^*) is weakly* convergent to x^* if and only if $x_\alpha^*x \to x^*x$ for each x in X, and (x_α^*) is weakly* Cauchy if and only if (x_α^*x) is a Cauchy (that is, convergent) net in $\mathbb{F}$ for each x in X. Notice also that by Proposition 2.4.12, a basis for the weak* topology of X^* is given by the collection of all subsets of X^* of the form

$$\{\, y^* : y^* \in X^*,\ |(y^* - x^*)x| < 1 \text{ for each } x \text{ in } A \,\}$$

such that $x^* \in X^*$ and A is a finite subset of X.

Since the weak* topology of the dual space of a normed space X is a subtopology of the norm topology of X^*, part (a) of the following result is a strengthening of part (a) of Proposition 1.10.15, while the rest of this result provides the characterizations of $({}^{\perp}B)^{\perp}$ for a subset B of X^* promised in the discussion following that proposition.

2.6.6 Proposition. *Let X be a normed space and let A and B be subsets of X and X^* respectively.*

(a) *The set $A^\perp$ is a weakly* closed subspace of X^*.*
(b) $({}^\perp B)^\perp = [B]^{w^*}$.
(c) *If B is a subspace of X^*, then $({}^\perp B)^\perp = \overline{B}^{w^*}$.*

PROOF. Let Q be the natural map from X into X^{**}. Then

$$A^\perp = \{\, x^* : x^* \in X^*,\ x^*x = 0 \text{ for each } x \text{ in } A \,\} = \bigcap \{\, \ker(Qx) : x \in A \,\}.$$

For each x in A, the linear functional Qx is weakly* continuous on X^*, from which it follows that $\bigcap\{\, \ker(Qx) : x \in A \,\}$ is a weakly* closed subspace of X^*, proving (a). For (b), first notice that $({}^\perp B)^\perp$ is a weakly* closed subspace of X^* that includes B, so $[B]^{w^*} \subseteq ({}^\perp B)^\perp$. Now suppose that $x_0^* \in X^* \setminus [B]^{w^*}$. An application of Corollary 2.2.20 then yields an x_0 in X such that $x_0^* x_0 = 1$ and $[B]^{w^*} \subseteq \ker(Qx_0)$. Since $x_0 \in {}^\perp B$, it follows that $x_0^* \notin ({}^\perp B)^\perp$. Therefore $({}^\perp B)^\perp \subseteq [B]^{w^*}$, which finishes the proof of (b). Since $\overline{\langle B \rangle}^{w^*} = [B]^{w^*}$ by Theorem 2.2.9 (i), part (c) follows immediately. ■

Suppose that X is a normed space and that $A \subseteq X^*$. In accordance with Definition 2.2.8, the set A is *weakly* bounded* if, for each weak* neighborhood U of 0 in X^*, there is a positive s_U such that $A \subseteq tU$ whenever $t > s_U$. By Proposition 2.4.14, the set A is weakly* bounded if and only if $\{\, x^*x : x^* \in A \,\}$ is bounded in $\mathbb{F}$ for each x in X. The following result is the weak* analog of the fact that a subset of a normed space is bounded if and only if it is weakly bounded. However, notice the requirement that X be a Banach space.

2.6.7 Theorem. *Let X be a Banach space. Then a subset of X^* is bounded if and only if it is weakly* bounded.*

PROOF. Since every weak* neighborhood of 0 in X^* is open with respect to the norm topology, it is clear that every bounded subset of X^* is weakly* bounded. Conversely, suppose that A is a weakly* bounded subset of X^*. It may be assumed that A is nonempty. Then $\sup\{\, |x^*x| : x^* \in A \,\}$ is finite for each x in the Banach space X, so the uniform boundedness principle implies that $\sup\{\, \|x^*\| : x^* \in A \,\}$ is finite, as required. ■

2.6.8 Corollary. *Let X be a Banach space. Then a subset A of X^* is bounded if and only if $\{\, x^*x : x^* \in A \,\}$ is a bounded set of scalars for each x in X.*

If X is a normed space, then it follows from Propositions 2.2.11, 2.2.12, and 2.1.47 that weakly* compact sets and weakly* Cauchy sequences in X^* are weakly* bounded and that weakly* convergent sequences in X^* are weakly* Cauchy. This immediately yields two more corollaries of Theorem 2.6.7.

2.6.9 Corollary. *Let X be a Banach space. Then weakly* compact subsets of X^* are bounded.*

2.6.10 Corollary. *Let X be a Banach space. Then weakly* Cauchy sequences in X^*, and so weakly* convergent sequences in X^*, are bounded.*

Neither the preceding theorem nor any of its three corollaries remain true if X is only required to be a normed space; see Exercise 2.68. However, the following result does hold for every normed space, whether or not it is complete.

2.6.11 Proposition. *Let X be an infinite-dimensional normed space. Then every nonempty weakly* open subset of X^* is unbounded.*

PROOF. Since X is infinite-dimensional, so is X^*. If A is a nonempty weakly* open subset of X^*, then A is also weakly open, and so is unbounded by Corollary 2.5.9. ■

The following two results are the weak* analogs of Propositions 2.5.14 and 2.5.15. The first of the two has a completeness hypothesis that cannot be omitted; see Exercise 2.70.

2.6.12 Proposition. *Let X be a Banach space. Then the weak* topology of X^* is induced by a metric if and only if X is finite-dimensional.*

PROOF. If X is finite-dimensional, then the weak* topology of X^* is induced by a metric since it is the same as the norm topology. For the converse, suppose that X is infinite-dimensional and that the weak* topology of X^* is induced by a metric d. For each positive integer n, let U_n be the d-open ball in X^* centered at 0 and having d-radius n^{-1}. By Proposition 2.6.11, each U_n contains some x_n^* such that $\|x_n^*\| \geq n$. It follows that (x_n^*) is an unbounded sequence in X^* that converges weakly* to 0, which contradicts Corollary 2.6.10. ■

2.6.13 Proposition. *Let X be a normed space. Then the weak* topology of X^* is complete if and only if X is finite-dimensional.*

PROOF. If X is finite-dimensional, then the weak* topology of X^* is the same vector topology as the norm topology, and therefore is complete by Corollary 2.1.51.

Suppose that X is infinite-dimensional. By Theorem 1.4.11, some linear functional f on X is unbounded. Let I be the collection of all finite subsets of X, preordered by declaring that $F_1 \preceq F_2$ when $F_1 \subseteq F_2$. For each F in I, obtain an x^*_F in X^* that agrees with f on F by first restricting f to the finite-dimensional subspace $\langle F\rangle$ of X, then letting x^*_F be any Hahn-Banach extension to X of this bounded linear functional on $\langle F\rangle$. The net $(x^*_F)_{F\in I}$ is weakly* Cauchy, since $x^*_F x = fx$ when $\{x\} \preceq F$, but cannot be weakly* convergent to any x^* in X^*, since that would require that $x^*x = \lim_F x^*_F x = fx$ for each x in X. ■

Just as norm functions are weakly lower semicontinuous, norm functions on dual spaces are weakly* lower semicontinuous.

2.6.14 Theorem. *Let X be a normed space. If (x^*_α) is a weakly* convergent net in X^*, then $\|w^*\text{-}\lim_\alpha x^*_\alpha\| \le \liminf_\alpha \|x^*_\alpha\|$.*

PROOF. Let $x^* = w^*\text{-}\lim_\alpha x^*_\alpha$. Suppose that $\epsilon > 0$. Then there is an x in B_X such that $\lim_\alpha |x^*_\alpha x| = |x^*x| \ge \|x^*\| - \epsilon$. Since there is an α_ϵ in the index set for (x^*_α) such that $\|x^*\| - 2\epsilon \le |x^*_\alpha x| \le \|x^*_\alpha\|$ when $\alpha_\epsilon \preceq \alpha$, it follows that $\|x^*\| \le \liminf_\alpha \|x^*_\alpha\|$. ■

The following partial converse of the preceding theorem will have an application in Example 5.4.13.

2.6.15 Theorem. *Suppose that X is a normed space, that $\|\cdot\|_a$ is a norm on X^* equivalent to its usual dual norm, and that $\|w^*\text{-}\lim_\alpha x^*_\alpha\|_a \le \liminf_\alpha \|x^*_\alpha\|_a$ whenever (x^*_α) is a weakly* convergent net in X^*. Then there is a norm $\|\cdot\|_b$ on X equivalent to its original norm such that $\|\cdot\|_a$ is the dual norm on $(X, \|\cdot\|_b)^*$.*

PROOF. Throughout this proof, the original norms of X and X^* will be denoted by $\|\cdot\|$. Let $\|x\|_b = \sup\{\,|x^*x| : x^* \in X^*,\ \|x^*\|_a \le 1\,\}$ whenever $x \in X$. It is easy to check that $\|\cdot\|_b$ is a norm on X. Now let s and t be positive reals such that $s\|x^*\| \le \|x^*\|_a \le t\|x^*\|$ whenever $x^* \in X^*$. If $x \in X$, then

$$\begin{aligned}\|x\|_b &= \sup\{\,|x^*x| : x^* \in X^*,\ \|x^*\|_a \le 1\,\}\\ &\le \sup\{\,|x^*x| : x^* \in X^*,\ s\|x^*\| \le 1\,\}\\ &= s^{-1}\|x\|\end{aligned}$$

and

$$\begin{aligned}\|x\| &= \sup\{\,|x^*x| : x^* \in X^*,\ \|x^*\| \le 1\,\}\\ &\le \sup\{\,|x^*x| : x^* \in X^*,\ t^{-1}\|x^*\|_a \le 1\,\}\\ &= t\|x\|_b,\end{aligned}$$

so $s\|x\|_b \le \|x\| \le t\|x\|_b$. Therefore $\|\cdot\|_b$ is equivalent to the original norm of X.

Let $\|\cdot\|_c$ be the dual norm on $(X, \|\cdot\|_b)^*$. All that remains to be proved is that $\|\cdot\|_a = \|\cdot\|_c$. Fix an x^* in X^*. If $\|x^*\|_a \le 1$, then $|x^*x| \le 1$ whenever $\|x\|_b \le 1$, so $\|x^*\|_c \le 1$. It follows that $\|x^*\|_c \le \|x^*\|_a$ no matter what the value of $\|x^*\|_a$ is. All that is left to be shown is that $\|x^*\|_c \ge \|x^*\|_a$.

Let ϵ be a fixed positive number. Suppose that $n \in \mathbb{N}$ and that $x_1, \dots, x_n$ is a basis for an n-dimensional subspace M_0 of X. Let $\big(T(x)\big)(x^*) = x^*x$ whenever $x \in X$ and $x^* \in X^*$. Then T is an isometric isomorphism from $(X, \|\cdot\|_b)$ into $(X^*, \|\cdot\|_a)^*$. If $\alpha_1, \dots, \alpha_n \in \mathbb{F}$, then

$$\begin{aligned}|\alpha_1 x^*x_1 + \cdots + \alpha_n x^*x_n| &\le \|x^*\|_c \, \|\alpha_1 x_1 + \cdots + \alpha_n x_n\|_b \\ &= \|x^*\|_c \, \|\alpha_1 Tx_1 + \cdots + \alpha_n Tx_n\|_{(X^*, \|\cdot\|_a)^*},\end{aligned}$$

so by Helly's theorem there is an $x^*_{M_0}$ in X^* such that

$$x^*_{M_0} x_j = (Tx_j)(x^*_{M_0}) = x^*x_j$$

when $j = 1, \dots, n$, and

$$\|x^*_{M_0}\|_a \le \|x^*\|_c + \epsilon.$$

Notice that $x^*_{M_0}$ and x^* agree on M_0. It follows that for each finite-dimensional subspace M of X there is an x^*_M in X^* such that x^*_M and x^* agree on M and $\|x^*_M\|_a \le \|x^*\|_c + \epsilon$. Preordering the collection of all finite-dimensional subspaces of X by declaring that $M_1 \preceq M_2$ when $M_1 \subseteq M_2$ makes (x^*_M) into a net. Since $x^*_M \xrightarrow{w^*} x^*$, it follows that

$$\|x^*\|_a \le \liminf_M \|x^*_M\|_a \le \|x^*\|_c + \epsilon.$$

Therefore $\|x^*\|_a \le \|x^*\|_c$ since ϵ is an arbitrary positive number. ■

So far, most of the results of this section have emphasized the similarities between the weak and weak* topologies, especially when the weak* topology is for the dual space of a Banach space. As the following two examples show, there are also some fundamental ways in which the two topologies differ.

2.6.16 Example. Let X be a nonreflexive Banach space and let x^{**} be any member of X^{**} that is not in the image of X under the natural map from X into X^{**}. Since x^{**} is continuous but not weakly* continuous, the kernel of x^{**} is a closed convex subset of X^* that is not weakly* closed. Thus, the weak* analog of Theorem 2.5.16 does not hold.

2.6.17 Example. Let c_0^* be identified with ℓ_1 in the usual way. Define $T\colon c_0^* \to c_0^*$ by the formula $T\big((\alpha_n)\big) = (\sum_n \alpha_n, \alpha_2, \alpha_3, \alpha_4, \dots)$. Then T is linear, is one-to-one since $\ker(T) = \{0\}$, and maps c_0^* onto itself since $T(\alpha_1 - \sum_{n=2}^{\infty} \alpha_n, \alpha_2, \alpha_3, \dots) = (\alpha_n)$ whenever $(\alpha_n) \in c_0^*$. Since $\|T(\alpha_n)\|_1 \le$

$2\|(\alpha_n)\|_1$ and $\|T^{-1}(\alpha_n)\|_1 \leq 2\|(\alpha_n)\|_1$ whenever $(\alpha_n) \in c_0^*$, the operator T is an isomorphism from c_0^* onto itself.

Now let (e_n^*) be the sequence of standard unit vectors of ℓ_1, viewed as members of c_0^*. Then (e_n^*) converges weakly* to 0, but the sequence (Te_n^*) does not since $(Te_n^*)(1,0,0,0,\dots) = 1$ for each n. The operator T is therefore not weak*-to-weak* continuous, even though it is a norm-to-norm isomorphism from c_0^* onto itself. Thus, the weak* analogs of Theorem 2.5.11 and Corollary 2.5.12 both fail.

As will be seen in Corollary 3.1.12, it is true that if X and Y are normed spaces and T is a weak*-to-weak* continuous linear operator from X^* into Y^*, then T is norm-to-norm continuous.

The next theorem illustrates another way in which the weak* topology is strikingly different from both the norm and the weak topology. As was mentioned in the introduction to this chapter, many of the difficulties that arise when trying to extend familiar facts about finite-dimensional normed spaces to the infinite-dimensional case come about because of the loss of the Heine-Borel property, that is, because the closed unit ball of a normed space X is not compact unless X is finite-dimensional. When X is infinite-dimensional, the fact that the weak topology of X is a proper subtopology of the norm topology makes it easier for B_X to be weakly compact than compact, and in fact B_X is weakly compact if and only if X is reflexive, as will be seen in Theorem 2.8.2. This leads naturally to the question of what conditions on X assure that B_{X^*} is weakly* compact.

One of the major results of the theory of normed spaces is that B_{X^*} is *always* weakly* compact. The general form of this result first appeared in a 1940 paper by Leonidas Alaoglu, and for that reason is often called Alaoglu's theorem. However, there is a result on page 123 of Banach's 1932 book [13] that easily implies the theorem when X is a separable Banach space; see Exercise 2.73 at the end of this section. For that reason, Banach is often given joint credit with Alaoglu for discovering the theorem, and that is what is done here.

2.6.18 The Banach-Alaoglu Theorem. (S. Banach, 1932 [13]; L. Alaoglu, 1940 [3]). *Let X be a normed space. Then B_{X^*} is weakly* compact.*

PROOF. Let Q be the natural map from X into X^{**}. Then $Q(B_X)$ is a separating topologizing family of functions on B_{X^*} that induces the restriction of the weak* topology of X^* to B_{X^*}. All references to a topology for B_{X^*} in the rest of this proof are to this topology. Let $I = \{\,\alpha : \alpha \in \mathbb{F},\ |\alpha| \leq 1\,\}$, let $I_x = I$ for each x in B_X, and let I^{B_X} be the topological product $\prod_{x\in B_X} I_x$. Then I^{B_X} is compact by Tychonoff's theorem, and the map $F\colon B_{X^*} \to I^{B_X}$ defined by the formula $F(x^*) = (x^*x)_{x\in B_X}$ is a homeomorphism from B_{X^*} onto a topological subspace of I^{B_X} by Proposition 2.4.7. The theorem will be proved once it is shown that $F(B_{X^*})$ is closed in I^{B_X}.

Let (x_β^*) be a net in B_{X^*} such that $F(x_\beta^*)$ converges to some $(\alpha_x)_{x\in B_X}$ in I^{B_X}. The goal is to find an x^* in B_{X^*} such that $F(x^*) = (\alpha_x)_{x\in B_X}$. For each nonzero x in X, let $x^*(x) = \|x\|\alpha_{(\|x\|^{-1}x)}$, and let $x^*(0) = 0$. Since $x^*(x) = \|x\| \lim_\beta x_\beta^*(\|x\|^{-1}x) = \lim_\beta x_\beta^* x$ whenever $x \in X \setminus \{0\}$, and thus $x^*(x) = \lim_\beta x_\beta^* x$ for every x in X, it follows that x^* is a linear functional on X, and furthermore that $x^* \in B_{X^*}$ since $|x^*x| = \lim_\beta |x_\beta^* x| \le \|x\|$ whenever $x \in X$. For each x in B_X,

$$\big(F(x^*)\big)_x = x^*x = \lim_\beta x_\beta^* x = \lim_\beta \big(F(x_\beta^*)\big)_x = \alpha_x,$$

so $F(x^*) = (\alpha_x)_{x\in B_X}$, as required. ∎

It is an obvious consequence of the Banach-Alaoglu theorem that every closed ball in the dual space of a normed space is weakly* compact, from which the next result follows immediately.

2.6.19 Corollary. *Let X be a normed space. Then every bounded subset of X^* is relatively weakly* compact. In particular, subsets of X^* that are bounded and weakly* closed are weakly* compact.*

2.6.20 Corollary. *Let X be a separable normed space and let A be a bounded subset of X^*. Then the relative weak* topology of A is induced by a metric.*

PROOF. It may be assumed that A is weakly* closed in X^* and therefore weakly* compact. Let D be a countable dense subset of X and let Q be the natural map from X into X^{**}. Since $Q(X)$ is a separating family of functions for A, so is the countable dense subset $Q(D)$ of $Q(X)$. It follows from Corollary 2.4.10 that the relative weak* topology of A is metrizable. ∎

2.6.21 Corollary. *Let X be a Banach space. Then every weakly* Cauchy sequence in X^* is weakly* convergent. That is, every Banach space has a weakly* sequentially complete dual space.*

PROOF. Let (x_n^*) be a weakly* Cauchy sequence in X^*. Corollary 2.6.10 implies that the sequence (x_n^*) is bounded, and therefore has a weakly* convergent subnet by the relative weak* compactness of bounded subsets of X^*. It follows from Proposition 2.1.46 that (x_n^*) is weakly* convergent to the limit of that subnet. ∎

Dual spaces of infinite-dimensional Banach spaces are therefore always weakly* sequentially complete but never weakly* complete. An example of an incomplete normed space whose dual space is not weakly* sequentially complete is given in Exercise 2.69.

2.6.22 Corollary. *Let X be a normed space. Then there is a compact Hausdorff space K such that X is isometrically isomorphic to a subspace* *of $C(K)$. If X is a Banach space, then X is isometrically isomorphic to a closed subspace of $C(K)$.*

PROOF. Let K be B_{X^*} with the topology it inherits from the weak* topology of X^*, and define $T\colon X \to C(K)$ by the formula $\big(T(x)\big)(x^*) = x^*x$. It is easy to check that T really does take its values in $C(K)$ and that T is linear. If $x \in X$, then it follows from Theorem 1.10.9 that

$$\|x\| = \max\{\,|x^*x| : x^* \in B_{X^*}\} = \max\{\,|(Tx)(x^*)| : x^* \in K\,\} = \|Tx\|_\infty,$$

so T is an isometric isomorphism from X into $C(K)$. If X is a Banach space, then so is $T(X)$, and therefore $T(X)$ is closed in $C(K)$. ■

Thus, in a sense the entire theory of normed spaces is contained in the theory of the subspaces of normed spaces $C(K)$ such that K is a compact Hausdorff space. This by no means trivializes the theory of normed spaces, but rather serves to point out the richness of the theory of the spaces $C(K)$.

Suppose that X is a normed space. If X is separable, then the relative weak* topology of B_{X^*} is metrizable by Corollary 2.6.20. It turns out that the converse is also true.

2.6.23 Theorem. *Let X be a normed space. Then the relative weak* topology of B_{X^*} is induced by a metric if and only if X is separable.*

PROOF. As was observed in the discussion preceding the theorem, one direction follows from Corollary 2.6.20. For the other, suppose that the relative weak* topology of B_{X^*} is induced by some metric d. For each positive integer n, let U_n be the d-open ball in B_{X^*} of radius n^{-1} centered at 0. By Proposition 2.4.12, there is for each n a finite subset A_n of X such that

$$\begin{aligned} U_n &\supseteq B_{X^*} \cap \{\, x^* : x^* \in X^*,\ |x^*x| < 1 \text{ for each } x \text{ in } A_n \,\} \\ &\supseteq B_{X^*} \cap \{\, x^* : x^* \in X^*,\ x^*x = 0 \text{ for each } x \text{ in } A_n \,\}, \end{aligned}$$

from which it follows that

$$\{0\} = \bigcap_n U_n = B_{X^*} \cap \Big\{\, x^* : x^* \in X^*,\ x^*x = 0 \text{ for each } x \text{ in } \bigcup_n A_n \Big\}.$$

Since $\{\, x^* : x^* \in X^*,\ x^*x = 0 \text{ for each } x \text{ in } \bigcup_n A_n \,\}$ is a subspace of X^*, it must be that

$$\Big\{\, x^* : x^* \in X^*,\ x^*x = 0 \text{ for each } x \text{ in } \bigcup_n A_n \Big\} = \{0\}.$$

Therefore by Proposition 1.10.15 (b),

$$\begin{aligned}\Big[\bigcup_n A_n\Big] &= {}^{\perp}\Big(\Big(\bigcup_n A_n\Big)^{\perp}\Big)\\ &= {}^{\perp}\Big\{x^* : x^* \in X^*,\, x^*x = 0 \text{ for each } x \text{ in } \bigcup_n A_n\Big\}\\ &= {}^{\perp}\{0\}\\ &= X.\end{aligned}$$

Since $\bigcup_n A_n$ is countable, it follows from Proposition 1.12.1 (a) that X is separable. ■

The second dual X^{**} of a normed space X is the dual of X^* and therefore has its own weak* topology. The rest of this section concerns that topology.

2.6.24 Proposition. *The natural map Q from a normed space X into X^{**} is weak-to-weak* continuous, and in fact is a weak-to-relative-weak* homeomorphism from X onto $Q(X)$.*

PROOF. Just notice that if (x_α) is a net in X and x is an element of X, then $x_\alpha \xrightarrow{w} x$ if and only if $x^*x_\alpha \to x^*x$ for each x^* in X^*, which happens if and only if $Qx_\alpha \xrightarrow{w^*} Qx$. ■

2.6.25 Corollary. *Let X be a normed space and let Q be the natural map from X into X^{**}. Then the topologies that $Q(X)$ inherits from the weak and weak* topologies of X^{**} are the same.*

PROOF. It follows from the preceding proposition and Corollary 2.5.12 that Q is both a weak-to-relative-weak* homeomorphism from X onto the subset $Q(X)$ of X^{**} and a weak-to-weak homeomorphism from X onto the normed space $Q(X)$. By Proposition 2.5.22, the weak topology of the normed space $Q(X)$ is the same as the relative weak topology of $Q(X)$ when $Q(X)$ is viewed as a subset of X^{**}, from which the corollary follows. ■

If Q is the natural map from a nonreflexive Banach space X into its second dual, then $Q(X)$ is a convex proper subset of X^{**} that is closed and therefore weakly closed, so $Q(X)$ is neither dense nor weakly dense in X^{**}. However, it turns out that $Q(X)$ must be weakly* dense in X^{**}, and in fact a bit more than that can be said.

2.6.26 Goldstine's Theorem. (H. H. Goldstine, 1938 [89]). *Let X be a normed space and let Q be the natural map from X into X^{**}. Then $Q(B_X)$ is weakly* dense in $B_{X^{**}}$.*

PROOF. It is enough to show that $B_{X^{**}} \subseteq \overline{Q(B_X)}^{w^*}$. Suppose that x_0^{**} is an element of X^{**} not in $\overline{Q(B_X)}^{w^*}$. It is enough to show that $\|x_0^{**}\| > 1$. Since

$\overline{Q(B_X)}^{w^*}$ is convex and weakly* closed, an application of Theorem 2.2.28 produces an x_0^* in X^* such that

$$\begin{aligned}|x_0^{**}x_0^*| &\geq \operatorname{Re} x_0^{**}x_0^* \\ &> \sup\{\operatorname{Re} x^{**}x_0^* : x^{**} \in \overline{Q(B_X)}^{w^*}\} \\ &\geq \sup\{\operatorname{Re} x_0^* x : x \in B_X\} \\ &= \|\operatorname{Re} x_0^*\| \\ &= \|x_0^*\|.\end{aligned}$$

It follows that $\|x_0^{**}\| > 1$. ■

2.6.27 Corollary. *Let X be a normed space and let Q be the natural map from X into X^{**}. Then $\overline{Q(B_X)}^{w^*} = B_{X^{**}}$.*

Proof. The Banach-Alaoglu theorem and Goldstine's theorem together imply that $Q(B_X)$ is weakly* dense in the weakly* closed set $B_{X^{**}}$, from which the corollary follows. ■

2.6.28 Corollary. *Let X be a normed space and let Q be the natural map from X into X^{**}. Then $Q(X)$ is weakly* dense in X^{**}.*

Proof. If x^{**} is a nonzero element of X^{**}, then Goldstine's theorem implies that there is a net (x_α) in B_X such that $Qx_\alpha \xrightarrow{w^*} \|x^{**}\|^{-1}x^{**}$, from which it follows that $Q(\|x^{**}\|x_\alpha) \xrightarrow{w^*} x^{**}$. ■

Exercises

2.67 Suppose that X is c_0 or ℓ_p, where $1 \leq p < \infty$, and that X^* is identified in the usual way with the appropriate ℓ_q such that $1 \leq q \leq \infty$. Let $((\beta_n^{(\alpha)}))_{\alpha \in I}$ be a net in X^* and (β_n) an element of X^*.

- (a) Show that if $(\beta_n^{(\alpha)}) \xrightarrow{w^*} (\beta_n)$, then $\beta_n^{(\alpha)} \to \beta_n$ for each n.
- (b) Show that if the net $((\beta_n^{(\alpha)}))_{\alpha \in I}$ is bounded and $\beta_n^{(\alpha)} \to \beta_n$ for each n, then $(\beta_n^{(\alpha)}) \xrightarrow{w^*} (\beta_n)$.
- (c) Show that the conclusion of (b) can fail if the net $((\beta_n^{(\alpha)}))_{\alpha \in I}$ is not required to be bounded.

2.68 Let X be the vector space of finitely nonzero sequences equipped with the ℓ_1 norm. For each positive integer m, let $x_m^* : X \to \mathbb{F}$ be defined by the formula $x_m^*((\alpha_n)) = m \cdot \alpha_m$. Let $A = \{x_m^* : m \in \mathbb{N}\}$.

- (a) Show that A is a weakly* bounded subset of X^* that is not norm bounded, and therefore that the conclusions of Theorem 2.6.7 and Corollary 2.6.8 do not follow when X is only required to be a normed space.

(b) Show that the conclusions of Corollaries 2.6.9 and 2.6.10 do not follow when X is only required to be a normed space.

2.69 Let X be the vector space of finitely nonzero sequences equipped with the ℓ_1 norm. Find a weakly* Cauchy sequence in X^* that is not weakly* convergent.

2.70 Give an example of an infinite-dimensional normed space X such that the weak* topology of X^* is metrizable.

2.71 Suppose that X is a normed space and that D is a dense subset of S_X.

(a) Show that a bounded net (x^*_α) in X^* converges weakly* to an element x^* of X^* if and only if $x^*_\alpha x \to x^* x$ for each x in D.

(b) Give an example to show that the requirement in (a) that the net be bounded cannot be omitted.

(c) Let $C[0,1]^*$ be identified with rca$[0,1]$ in the usual way. Show that a bounded net (μ_α) in $C[0,1]^*$ converges weakly* to an element μ of $C[0,1]^*$ if and only if $\int_{[0,1]} t^n \, d\mu_\alpha(t) \to \int_{[0,1]} t^n \, d\mu(t)$ for each nonnegative integer n.

2.72 Suppose that X is a separable normed space. The goal of this exercise is to show that there is a norm $\|\cdot\|_0$ on X^* such that for each bounded subset A of X^*, the topologies induced on A by the weak* and $\|\cdot\|_0$ topologies of X^* are the same. It may be assumed that $X \neq \{0\}$. Let $\{ x_n : n \in \mathbb{N} \}$ be a countable dense subset of S_X.

(a) Define $\|\cdot\|_0 \colon X^* \to \mathbb{R}$ by the formula $\|x^*\|_0 = \sum_n 2^{-n}|x^* x_n|$. Show that $\|\cdot\|_0$ is a norm on X^*, and that if $\|\cdot\|$ is the usual norm of X^*, then $\|x^*\|_0 \leq \|x^*\|$ for each x^* in X^*.

(b) Let A be a bounded subset of X^*. Show that the topologies that A inherits from the weak* and $\|\cdot\|_0$ topologies of X^* are the same. Exercise 2.71 might help.

Do not use the Banach-Alaoglu theorem or any of its corollaries in your arguments. Notice that this provides a proof of Corollary 2.6.20 that is not based on the Banach-Alaoglu theorem.

2.73 The following result appears on page 123 of Banach's book [13], using slightly different notation and terminology: *If the Banach space X is separable, then every bounded sequence in X^* has a weakly* convergent subsequence.* Prove that this implies the conclusion of the Banach-Alaoglu theorem when X is a separable Banach space. Of course, you should not use the Banach-Alaoglu theorem or any of its corollaries in your argument. You might find Exercise 2.72 helpful.

2.74 For this exercise, it will be useful to know that if K is a compact metric space and P is the Cantor set, then there is a continuous map from P onto K. (I. Rosenholtz has given a particularly nice proof of this that is built up from a sequence of elementary lemmas; see [198].) Let X be a separable normed space.

(a) Prove that X is isometrically isomorphic to a subspace of $C(P)$.

(b) Prove that X is isometrically isomorphic to a subspace of $C[0,1]$.

2.75 (a) Prove that if X is a separable Banach space, then every weakly* compact subset of X^* is weakly* sequentially compact.

(b) Find a subset of ℓ_∞^* that is weakly* compact but not weakly* sequentially compact.

2.76 Use the Banach-Alaoglu theorem to prove that every bounded net in a normed space has a weakly Cauchy subnet. (See Exercise 2.66 for another line of proof that uses ultranets.)

2.77 Corollary 2.6.28 can be easily derived from results of this section much more basic than Goldstine's theorem. Do so.

2.78 The purpose of this exercise is to generalize the notion of a weak* topology to a more abstract setting. Let X be a topological vector space and let Y be the subset of $(X^*)^\#$ consisting of all linear functionals f on X^* for which there is an x_f in X such that $fx^* = x^*x_f$ for each x^* in X^*. Then the *weak* topology of* X^* is the topology induced on X^* by the topologizing family Y.

(a) Prove that Y is a subspace of $(X^*)^\#$ that is a separating family of functions for X^*, and therefore that the weak* topology of X^* is a completely regular locally convex topology for X^* such that the dual space of X^* with respect to this topology is Y.

Suppose that $A \subseteq X$. Then $\{ x^* : x^* \in X^*, |x^*x| \leq 1 \text{ for each } x \text{ in } A \}$ is called the *absolute polar set for* A.

(b) Let U be a neighborhood of 0 in X. Prove that the absolute polar set for U is weakly* compact.

(c) Derive the Banach-Alaoglu theorem from (b).

The statement proved in (b) is itself sometimes called the Banach-Alaoglu theorem.

2.7 The Bounded Weak* Topology

Most of this section is devoted to the study of yet another locally convex topology for the dual space of a normed space, the bounded weak* topology. Though this topology is interesting in its own right, the primary reason for introducing it is to obtain a theorem due to M. G. Krein (pronounced "crane") and V. L. Šmulian that provides a useful way to test whether a convex subset of the dual space of a Banach space is weakly* closed.

2.7.1 Definition. (J. Dieudonné, 1950 [63]). Let X be a normed space. For each x^* in X^* and each sequence (x_n) in X that converges to 0, let

$$B\big(x^*, (x_n)\big) = \{ y^* : y^* \in X^*, |(y^* - x^*)x_n| < 1 \text{ for each } n \}.$$

The topology for X^* having the basis consisting of all such sets $B\big(x^*, (x_n)\big)$ is the *bounded weak* topology of* X^*.

Of course, it must be shown that the collection of sets $B(x^*,(x_n))$ really is a basis for a topology for X^*. This is done in the proof of the next theorem.

For notational convenience, properties related to the bounded weak* topology will often be referred to as "b-weak*" properties; for example, a set that is closed with respect to the bounded weak* topology will be said to be b-weakly* closed. Convergence of a net (x^*_α) to an x^* with respect to this topology will be denoted by writing $x^*_\alpha \xrightarrow{bw^*} x^*$.

2.7.2 Theorem. *Let X be a normed space. Then the bounded weak* topology of X^* is a completely regular locally convex topology. Let $\mathfrak{T}_{bw^*}$ be this topology, and let $\mathfrak{T}_{w^*}$ and $\mathfrak{T}_n$ be, respectively, the weak* and norm topologies of X^*. Then $\mathfrak{T}_{w^*} \subseteq \mathfrak{T}_{bw^*} \subseteq \mathfrak{T}_n$. If A is a bounded subset of X^*, then the relative bounded weak* and relative weak* topologies of A are the same. If $\mathfrak{T}$ is a topology for X^* such that the relative topologies inherited by each bounded subset of X^* from $\mathfrak{T}_{w^*}$ and $\mathfrak{T}$ are the same, then $\mathfrak{T} \subseteq \mathfrak{T}_{bw^*}$.*

PROOF. Let $\mathfrak{B}$ be the collection of all subsets $B(x^*,(x_n))$ of X^* as in Definition 2.7.1. It is necessary to check that $\mathfrak{B}$ really is a basis for a topology for X^*. Since $X^* = B(0,(0,0,\dots))$, it follows trivially that $X^* = \bigcup\{\, B : B \in \mathfrak{B}\,\}$. Now suppose that $B(x_1^*,(x_{1,n})), B(x_2^*,(x_{2,n})) \in \mathfrak{B}$ and that $x_0^* \in B(x_1^*,(x_{1,n})) \cap B(x_2^*,(x_{2,n}))$. Since $x_{1,n} \to 0$ and $x_{2,n} \to 0$, it follows that $\{\, |(x_0^* - x_j^*)x_{j,n}| : j \in \{1,2\},\ n \in \mathbb{N}\,\}$ has a maximum element less than 1. Let δ be such that

$$0 < \delta < 1 - \max\{\, |(x_0^* - x_j^*)x_{j,n}| : j \in \{1,2\},\ n \in \mathbb{N}\,\},$$

and let $(x_{0,n})$ be the sequence

$$(\delta^{-1}x_{1,1},\ \delta^{-1}x_{2,1},\ \delta^{-1}x_{1,2},\ \delta^{-1}x_{2,2},\ \delta^{-1}x_{1,3},\ \delta^{-1}x_{2,3},\ \dots),$$

a sequence in X that converges to 0. If $x^* \in B(x_0^*,(x_{0,n}))$, then

$$|(x^* - x_j^*)x_{j,n}| \le |(x^* - x_0^*)x_{j,n}| + |(x_0^* - x_j^*)x_{j,n}| < \delta + (1-\delta) = 1$$

whenever $n \in \mathbb{N}$ and j is 1 or 2, so $x^* \in B(x_1^*,(x_{1,n})) \cap B(x_2^*,(x_{2,n}))$. It follows that $x_0^* \in B(x_0^*,(x_{0,n})) \subseteq B(x_1^*,(x_{1,n})) \cap B(x_2^*,(x_{2,n}))$, which finishes the proof that $\mathfrak{B}$ is a basis for a topology for X^*.

The argument just given yields another fact that will be useful later in this proof. Suppose that $x^* \in X^*$, that (x_n) is a sequence in X converging to 0, and that $x_0^* \in B(x^*,(x_n))$. Letting $x_1^* = x_2^* = x^*$ and $(x_{1,n}) = (x_{2,n}) = (x_n)$ in the argument of the preceding paragraph shows that there is a sequence $(x_{0,n})$ in X converging to 0 such that $x_0^* \in B(x_0^*,(x_{0,n})) \subseteq B(x^*,(x_n))$.

To see that $\mathfrak{T}_{w^*} \subseteq \mathfrak{T}_{bw^*} \subseteq \mathfrak{T}_n$, first observe that for each x^* in X^* and each x in X,

$$B(x^*,(x,0,0,\dots)) = \{\, y^* : y^* \in X^*,\ |(y^* - x^*)x| < 1\,\}.$$

It follows from this and Proposition 2.4.12 that

$$\{ B(x^*, (x, 0, 0, \dots)) : x^* \in X^*, x \in X \}$$

is a subbasis for $\mathfrak{T}_{w^*}$, and therefore that $\mathfrak{T}_{w^*} \subseteq \mathfrak{T}_{bw^*}$. Notice that this implies that $\mathfrak{T}_{bw^*}$ is Hausdorff. Now suppose that $(x^*_\alpha)_{\alpha \in I}$ is a net in X^* that is norm convergent to an x^* in X^*. Then for each sequence (x_n) in X converging to 0 there is an $\alpha_{(x_n)}$ in I such that $x^*_\alpha \in B(x^*, (x_n))$ when $\alpha_{(x_n)} \preceq \alpha$. This implies that $x^*_\alpha \xrightarrow{bw^*} x^*$, so $\mathfrak{T}_{bw^*} \subseteq \mathfrak{T}_n$.

The next order of business is to show that $\mathfrak{T}_{bw^*}$ is locally convex. Suppose that $x^*, y^* \in X^*$, that $\alpha \in \mathbb{F}$, and that U and V are b-weak* neighborhoods of $x^* + y^*$ and αx^* respectively. Then there are sequences (u_n) and (v_n) in X converging to 0 such that $x^* + y^* \in B(x^* + y^*, (u_n)) \subseteq U$ and $\alpha x^* \in B(\alpha x^*, (v_n)) \subseteq V$. It is easy to check that

$$B(x^*, (2u_n)) + B(y^*, (2u_n)) \subseteq B(x^* + y^*, (u_n)),$$

from which it follows that addition of vectors is $\mathfrak{T}_{bw^*}$-continuous. If

$$|\beta - \alpha| < \min \left\{ \frac{1}{2(|x^* v_n| + 1)} : n \in \mathbb{N} \right\}$$

and $w^* \in B(x^*, ((2|\alpha| + 1)v_n))$, then a straightforward computation shows that $|(\beta w^* - \beta x^*)v_n| < \frac{1}{2}$ and $|(\beta x^* - \alpha x^*)v_n| < \frac{1}{2}$ for each n, so $\beta w^* \in B(\alpha x^*, (v_n))$. Multiplication of vectors by scalars is therefore $\mathfrak{T}_{bw^*}$-continuous, so $\mathfrak{T}_{bw^*}$ is a vector topology. It then follows that $\mathfrak{T}_{bw^*}$ is locally convex since its basis $\mathfrak{B}$ consists of convex sets. Because this vector topology is Hausdorff, it is completely regular by Theorem 2.2.14.

Now suppose that A is a bounded subset of X^*. Let x^* be a member of X^* and let (x_n) be a sequence in X converging to 0. Then there is a positive integer n_0 such that $|(y^* - x^*)x_n| < 1$ whenever $y^* \in A$ and $n > n_0$. It follows that

$$\begin{aligned} &A \cap B(x^*, (x_n)) \\ &\qquad = A \cap \{ y^* : y^* \in X^*, |(y^* - x^*)x_n| < 1 \text{ when } n = 1, \dots, n_0 \}. \end{aligned}$$

Since $A \cap \{ y^* : y^* \in X^*, |(y^* - x^*)x_n| < 1 \text{ when } n = 1, \dots, n_0 \}$ is open with respect to the relative topology that A inherits from $\mathfrak{T}_{w^*}$, it is clear that the relative topologies that A inherits from $\mathfrak{T}_{bw^*}$ and $\mathfrak{T}_{w^*}$ are the same.

Let $\mathfrak{T}$ be a topology for X^* such that the relative topologies inherited by each bounded subset of X^* from $\mathfrak{T}_{w^*}$ and $\mathfrak{T}$ are the same, and let U be a member of $\mathfrak{T}$. The proof of the theorem will be finished once it is shown that $U \in \mathfrak{T}_{bw^*}$. Let x^* be an element of U. The proof that U is b-weakly* open is based on the construction of a sequence $F_0, F_1, F_2, \dots$ of finite subsets of X such that for each nonnegative integer n,

(1) $\{ y^* : y^* \in X^*, \|y^* - x^*\| \le n+1, |(y^* - x^*)x| \le 1$ when $x \in F_0 \cup \cdots \cup F_n \} \subseteq U$; and

(2) $\|x\| \le n^{-1}$ if $n \ge 1$ and $x \in F_n$.

The construction begins with the observation that since the closed ball of radius 1 centered at x^* inherits the same relative topologies from $\mathfrak{T}$ and $\mathfrak{T}_{w^*}$, there are elements $x_1, \ldots, x_{n_0}$ of X such that

$$\{ y^* : y^* \in X^*, \|y^* - x^*\| \le 1, |(y^* - x^*)x_j| < 1 \text{ when } j = 1, \ldots, n_0 \} \subseteq U.$$

Let $F_0 = \{2x_1, \ldots, 2x_{n_0}\}$. Now suppose that $m \in \mathbb{N}$ and that finite subsets $F_0, \ldots, F_{m-1}$ of X have been found such that (1) and (2) are satisfied when $n = 0, \ldots, m-1$. Suppose further that there is no finite subset F_m of X such that (1) and (2) are satisfied when $n = m$. Let I be the collection of all finite subsets of $m^{-1}B_X$, preordered by declaring that $F \preceq G$ when $F \subseteq G$. For each F in I, let

$$K(F) = \{ y^* : y^* \in X^*, \|y^* - x^*\| \le m+1,$$
$$|(y^* - x^*)x| \le 1 \text{ when } x \in F_0 \cup \cdots \cup F_{m-1} \cup F \} \setminus U.$$

Notice that each $K(F)$ is nonempty and that $K(F) \supseteq K(G)$ when $F \preceq G$. As a straightforward consequence of the Banach-Alaoglu theorem and the fact that $\mathfrak{T}$ and $\mathfrak{T}_{w^*}$ induce the same relative topology on bounded subsets of X^*, each $K(F)$ is weakly* compact. For each F in I, let x_F^* be an element of $K(F)$. Then $(x_F^*)_{F \in I}$ is a net in the weakly* compact set $K(\varnothing)$ and has the property that $x_G^* \in K(F)$ whenever $F \preceq G$, and so has a subnet with a weak* limit x_0^* in $\bigcap \{ K(F) : F \in I \}$. It follows that $|(x_0^* - x^*)x| \le 1$ whenever $x \in m^{-1}B_X$, so $\|x_0^* - x^*\| \le m$. This means that

$$x_0^* \in \{ y^* : y^* \in X^*, \|y^* - x^*\| \le m,$$
$$|(y^* - x^*)x| \le 1 \text{ when } x \in F_0 \cup \cdots \cup F_{m-1} \},$$

but this last set lies inside U by (1) when $n = m-1$. This is a contradiction, since $x_0^* \in K(\varnothing) \subseteq X^* \setminus U$. It follows that there is a finite subset F_m of X such that (1) and (2) are satisfied when $n = m$, which means that the construction of the sequence $F_0, F_1, F_2, \ldots$ can be accomplished. It may be assumed that each F_n is nonempty, since each empty F_n may be replaced by $\{0\}$.

By (2), the elements of $\bigcup_{n=0}^{\infty} F_n$ can be listed as a sequence (x_n) that converges to 0. It follows from (1) that $B(x^*, (x_n)) \subseteq U$, which is enough to establish that U is b-weakly* open. ■

A useful fact obtained in the proof of the preceding theorem is worth repeating. With all notation as in Definition 2.7.1, suppose that x^* is an element of a b-weakly* open subset U of X^*. Then there is a sequence (x_n)

in X converging to 0 such that $B(x^*, (x_n)) \subseteq U$. That is, there is a basic b-weak* neighborhood of x^* "centered" at x^* that is included in U.

Suppose that (x^*_α) is a bounded net in the dual space of a normed space and that x^* is an element of that dual space. If A is the set consisting of x^* and the terms of (x^*_α), then the preceding theorem assures that the relative bounded weak* and relative weak* topologies of A are the same. The following result is an immediate consequence.

2.7.3 Corollary. *Let X be a normed space, let (x^*_α) be a bounded net in X^*, and let x^* be an element of X^*. Then $x^*_\alpha \xrightarrow{bw^*} x^*$ if and only if $x^*_\alpha \xrightarrow{w^*} x^*$.*

The characterizations of b-weakly* open sets in the following corollary are often used to define the bounded weak* topology.

2.7.4 Corollary. (J. Dieudonné, 1950 [63]). *Let X be a normed space and let A be a subset of X^*. Then the following are equivalent.*

(a) *The set A is b-weakly* open.*

(b) *The set $A \cap B$ is relatively weakly* open in B whenever B is a bounded subset of X^*.*

(c) *The set $A \cap tB_{X^*}$ is relatively weakly* open in tB_{X^*} whenever $t > 0$.*

PROOF. Let $\mathfrak{T}$ be the collection of all subsets U of X^* such that $U \cap B$ is relatively weakly* open in B whenever B is a bounded subset of X^*. It is easy to check that $\mathfrak{T}$ is a topology for X^*, that $\mathfrak{T}$ and the weak* topology of X^* induce the same relative topology on each bounded subset of X^*, and that $\mathfrak{T}' \subseteq \mathfrak{T}$ whenever $\mathfrak{T}'$ is a topology for X^* such that $\mathfrak{T}'$ and the weak* topology of X^* induce the same relative topology on each bounded subset of X^*. By the preceding theorem, the topology $\mathfrak{T}$ is the bounded weak* topology of X^*, which shows that (a) and (b) are equivalent. It is clear that (b) $\Rightarrow$ (c). Also, for each bounded subset B of X^* there is a positive t_B such that $B \subseteq t_B B_{X^*}$, from which it easily follows that (c) $\Rightarrow$ (b). ■

2.7.5 Corollary. (J. Dieudonné, 1950 [63]). *Let X be a normed space and let A be a subset of X^*. Then the following are equivalent.*

(a) *The set A is b-weakly* closed.*

(b) *The set $A \cap B$ is relatively weakly* closed in B whenever B is a bounded subset of X^*.*

(c) *The set $A \cap tB_{X^*}$ is weakly* closed in X^* whenever $t > 0$.*

(d) *The set A contains the weak* limits of all bounded nets in A that are weakly* convergent in X^*.*

PROOF. The equivalence of (a), (b), and (c) follows from the preceding corollary and the fact that tB_{X^*} is weakly* compact, hence weakly* closed, whenever $t > 0$. The equivalence of (c) and (d) is clear. ■

One pleasant property of the bounded weak* topology is that in contrast to what happens with the weak* topology, completeness is not lost when passing from the finite-dimensional to the infinite-dimensional case.

2.7.6 Proposition. *Let X be a normed space. Then the bounded weak* topology of X^* is complete.*

PROOF. Let $(x_\alpha^*)_{\alpha\in I}$ be a net in X^* that is b-weakly* Cauchy. Then (x_α^*) is also weakly* Cauchy, so $(x_\alpha^* x)$ converges for each x in X. Define $f\colon X \to \mathbb{F}$ by the formula $f(x) = \lim_\alpha x_\alpha^* x$. Then f is a linear functional on X. It will be shown that $f \in X^*$ and that $x_\alpha^* \xrightarrow{bw^*} f$.

Let (x_n) be a sequence in X converging to 0. For each positive ϵ, the set $\{\, y^* : y^* \in X^*,\ |y^* x_n| < \epsilon \text{ for each } n \,\}$ is a b-weak* neighborhood of 0; in the notation of Definition 2.7.1, it is the set $B\big(0, (\epsilon^{-1}x_n)\big)$. It follows from this and the fact that (x_α^*) is b-weakly* Cauchy that the net $\big((x_\alpha^* x_n)\big)_{\alpha\in I}$ in c_0 is Cauchy, hence convergent, and therefore has limit (fx_n). Since $(fx_n) \in c_0$, the linear functional f must be bounded, for otherwise the sequence (x_n) could have been chosen so that $|fx_n| \to +\infty$ even though $x_n \to 0$. The convergence of the net $\big(((x_\alpha^* - f)x_n)\big)_{\alpha\in I}$ to 0 in c_0 assures that there is an α_0 such that $x_\alpha^* \in B\big(f, (x_n)\big)$ when $\alpha_0 \preceq \alpha$. Since each b-weak* neighborhood of f includes a b-weak* neighborhood of f of the form $B\big(f, (y_n)\big)$, where (y_n) is a sequence in X converging to 0, it follows that $x_\alpha^* \xrightarrow{bw^*} f$. ■

2.7.7 Corollary. *Let X be an infinite-dimensional normed space. Then the bounded weak* topology of X^* is different from the weak*, weak, and norm topologies of X^*.*

PROOF. Since the bounded weak* topology of X^* is complete and the weak* and weak topologies of X^* are not, the bounded weak* topology is not equal to either of those. The bounded weak* and norm topologies of X^* must also differ, since B_{X^*} is b-weakly* compact but not norm compact. ■

It follows from the preceding corollary that if X is an infinite-dimensional normed space, then some subset of X^* is b-weakly* closed but not weakly* closed, and therefore is not weakly* closed even though it contains the weak* limits of all the bounded nets in the set that are weakly* convergent in X^*.

As was suggested in the introduction to this section, much can be learned about the weak* topology of the dual space X^* of a normed space X by studying the bounded weak* topology of X^*. The main reason for this is contained in the next theorem: The dual spaces with respect to the two topologies are the same if X is a Banach space. When interpreting this theorem and its corollaries, keep in mind that a linear functional on X^* is weakly* continuous if and only if it has the form $x^* \mapsto x^*x_0$ for some x_0 in X.

2.7.8 Theorem. *Let X be a Banach space. Then a linear functional on X^* is weakly* continuous if and only if it is b-weakly* continuous. That is, the dual spaces of X^* with respect to the weak* and bounded weak* topologies are the same.*

PROOF. Since every weakly* open subset of X^* is b-weakly* open, every weakly* continuous linear functional on X^* is b-weakly* continuous. For the converse, suppose that f is a b-weakly* continuous linear functional on X^*. A moment's thought about basic b-weak* neighborhoods of 0 "centered" at 0 shows that there is a sequence (x_n) in X converging to 0 such that $|fx^*| < 1$ whenever $x^* \in X^*$ and $|x^*x_n| < 1$ for each n. Notice that this implies that $fx^* = 0$ when $x^*x_n = 0$ for each n.

Define $T\colon X^* \to c_0$ by the formula $T(x^*) = (x^*x_n)$. Then T is clearly linear and bounded. It is also easy to check that the map $v^*: T(X^*) \to \mathbb{F}$ given by the formula $v^*(Tx^*) = fx^*$ is well-defined, linear, and bounded. By the normed space version of the Hahn-Banach extension theorem, there is a w^* in c_0^* whose restriction to $T(X^*)$ is v^*. Let (α_n) be the element of ℓ_1 that represents w^* in the usual way. For each x^* in X^*,

$$fx^* = v^*(Tx^*) = w^*(x^*x_n) = \sum_n \alpha_n x^*x_n = x^*\Big(\sum_n \alpha_n x_n\Big),$$

where the convergence of the last sum is assured by the completeness of X. It follows that f is in the image of X in X^{**} under the natural map, and so is weakly* continuous. ■

2.7.9 Corollary. *Let f be a linear functional on the dual space of a Banach space X. Then f is weakly* continuous if and only if the restriction of f to B_{X^*} is continuous with respect to the relative weak* topology of B_{X^*}.*

PROOF. If f is weakly* continuous, then its restriction to B_{X^*} is certainly continuous with respect to the relative weak* topology of B_{X^*}. Conversely, suppose that the restriction of f to B_{X^*} is continuous with respect to the relative weak* topology of B_{X^*}. Since B_{X^*} is weakly* closed in X^*, so is $\ker f \cap B_{X^*}$. If $t > 0$, then $\ker f \cap tB_{X^*} = t(\ker f \cap B_{X^*})$, which is weakly* closed, so $\ker f$ is b-weakly* closed by Corollary 2.7.5. It follows from Theorem 2.2.16 that f is b-weakly* continuous, and therefore weakly* continuous. ■

2.7.10 Corollary. *A linear functional on the dual space of a separable Banach space is weakly* continuous if and only if it is weakly* sequentially continuous.*

PROOF. Let X be a separable Banach space and let f be a weakly* sequentially continuous linear functional on X^*. Since continuity always implies sequential continuity, the corollary will be proved once it is shown that f is weakly* continuous. By the preceding corollary, it is enough to show that the restriction of f to B_{X^*} is continuous with respect to the relative weak* topology of B_{X^*}. Since this relative topology is metrizable by Corollary 2.6.20, the continuity of f on B_{X^*} with respect to this topology is assured by the weak* sequential continuity of f. ■

Theorem 2.7.8 has the main result of this section as an easy corollary.

2.7.11 The Krein-Šmulian Theorem on Weakly* Closed Convex Sets. (M. G. Krein and V. L. Šmulian, 1940 [146]). *Let C be a convex subset of the dual space of a Banach space X. Then C is weakly* closed if and only if $C \cap tB_{X^*}$ is weakly* closed whenever $t > 0$.*

PROOF. If C is weakly* closed and $t > 0$, then the weak* compactness of tB_{X^*} assures that $C \cap tB_{X^*}$ is weakly* closed. Conversely, suppose that $C \cap tB_{X^*}$ is weakly* closed whenever $t > 0$. Then C is b-weakly* closed by Corollary 2.7.5. Since the dual spaces of X^* with respect to the weak* and bounded weak* topologies are the same, it follows from Corollary 2.2.29 that C is weakly* closed. ■

2.7.12 Corollary. *Let M be a subspace of the dual space of a Banach space X. Then M is weakly* closed if and only if $M \cap B_{X^*}$ is weakly* closed.*

PROOF. Since $M \cap tB_{X^*} = t(M \cap B_{X^*})$ whenever $t > 0$, it follows that $M \cap tB_{X^*}$ is weakly* closed for each positive t if and only if $M \cap B_{X^*}$ is weakly* closed, which together with an application of the preceding theorem finishes the proof. ■

2.7.13 Corollary. *Let X be a separable Banach space. Then a convex subset of X^* is weakly* closed if and only if it is weakly* sequentially closed. In particular, every weakly* sequentially closed subspace of X^* is weakly* closed.*

PROOF. Suppose that C is a weakly* sequentially closed convex subset of X^*. It is enough to show that C is weakly* closed. Let t be a positive number. It is enough to show that $C \cap tB_{X^*}$ is weakly* closed. By Corollary 2.6.20, the relative weak* topology of the weakly* closed set tB_{X^*} is metrizable, so it is enough to show that $C \cap tB_{X^*}$ is weakly* sequentially closed. This follows immediately from the fact that both C and tB_{X^*} are weakly* sequentially closed. ■

Exercises

2.79 Prove that a bounded net in the dual space of a normed space is b-weakly* Cauchy if and only if it is weakly* Cauchy.

2.80 Let X be an infinite-dimensional normed space. Prove that every nonempty b-weakly* open subset of X^* is unbounded.

2.81 Let X be a Banach space. Prove that a subset of X^* is bounded if and only if it is b-weakly* bounded.

2.82 Let X be a Banach space. Prove that the bounded weak* topology of X^* is metrizable if and only if X is finite-dimensional. Exercise 2.80 might be helpful.

2.83 The purpose of this exercise is to show that if X is a Banach space and $\mathfrak{T}$ is a topology for X^* such that the relative topologies inherited by each bounded subset of X^* from $\mathfrak{T}$ and the weak* topology $\mathfrak{T}_{w^*}$ are the same, then it does *not* necessarily follow that $\mathfrak{T}_{w^*} \subseteq \mathfrak{T}$, even if $\mathfrak{T}$ is a completely regular locally convex topology. To this end, let $X = c_0$, let Q be the natural map from c_0 into c_0^{**}, let Y be the subspace of c_0 consisting of the finitely nonzero sequences, and let $\mathfrak{T}$ be the $Q(Y)$ topology of c_0^*.

(a) Prove that $\mathfrak{T}$ is a completely regular locally convex topology for c_0^* such that $\mathfrak{T} \subsetneqq \mathfrak{T}_{w^*}$.

(b) Prove that the relative topologies inherited by each bounded subset of c_0^* from $\mathfrak{T}$ and $\mathfrak{T}_{w^*}$ are the same. Exercise 2.67 might be helpful.

2.84 Let C be a convex subset of a normed space X.

(a) Prove that C is closed if and only if $C \cap tB_X$ is closed whenever $t > 0$.

(b) Prove that C is weakly closed if and only if $C \cap tB_X$ is weakly closed whenever $t > 0$.

These are, of course, the analogs for the norm and weak topologies of the Krein-Šmulian theorem on weakly* closed convex sets. Notice that these analogs do not require X to be complete.

2.85 The hypotheses of both Theorem 2.7.8 and the Krein-Šmulian theorem on weakly* closed convex sets include the requirement that the normed space X in question be a Banach space. The purpose of this exercise is to show that this requirement cannot in general be omitted. Let X be an incomplete normed space; let Q be the natural map from X into X^{**}; let Y be the closure of $Q(X)$ in X^{**}; let $\mathfrak{T}_{w^*}$, $\mathfrak{T}_{bw^*}$, and $\mathfrak{T}_Y$ be, respectively, the weak*, bounded weak*, and Y topologies of X^*; and let $(X^*_{w^*})^*$, $(X^*_{bw^*})^*$, and $(X^*_Y)^*$ be the respective dual spaces of X^* under these topologies.

(a) Show that the relative topologies inherited by each bounded subset of X^* from $\mathfrak{T}_{w^*}$ and $\mathfrak{T}_Y$ are the same.

(b) Show that $(X^*_{w^*})^* \subsetneqq (X^*_Y)^* \subseteq (X^*_{bw^*})^*$. Conclude that the requirement in the hypotheses of Theorem 2.7.8 that the normed space be complete is necessary.

(c) Find a convex subset C of X^* that is not weakly* closed even though $C \cap tB_{X^*}$ is weakly* closed whenever $t > 0$.

2.86 *The bounded weak topology.* Let X be a normed space and let $\mathfrak{T}_{bw}$ be the collection of all subsets U of X such that $U \cap B$ is relatively weakly open in B whenever B is a bounded subset of X. Then $\mathfrak{T}_{bw}$ is the *bounded weak topology of* X. For convenience, properties related to this topology will be called "b-weak" properties. Let $\mathfrak{T}_w$ and $\mathfrak{T}_n$ be the weak and norm topologies of X respectively. Prove each of the following statements.

(a) The collection $\mathfrak{T}_{bw}$ is a topology for X.

(b) $\mathfrak{T}_w \subseteq \mathfrak{T}_{bw} \subseteq \mathfrak{T}_n$.

(c) The relative topologies inherited by each bounded subset of X from $\mathfrak{T}_w$ and $\mathfrak{T}_{bw}$ are the same. Furthermore, if $\mathfrak{T}$ is a topology for X such that the relative topologies inherited by each bounded subset of X from $\mathfrak{T}_w$ and $\mathfrak{T}$ are the same, then $\mathfrak{T} \subseteq \mathfrak{T}_{bw}$.

(d) A bounded net (x_α) in X is b-weakly convergent to some x in X if and only if (x_α) is weakly convergent to x.

(e) A subset A of X is b-weakly open if and only if $A \cap tB_X$ is relatively weakly open in tB_X whenever $t > 0$.

(f) A subset A of X is b-weakly closed if and only if $A \cap B$ is relatively weakly closed in B whenever B is a bounded subset of X, which happens if and only if $A \cap tB_X$ is weakly closed in X whenever $t > 0$.

2.87 Let X be a normed space. The purpose of this exercise is to study another topology for X related to the bounded weak topology developed in Exercise 2.86. For each x in X and each sequence (x_n^*) in X^* that converges to 0, let

$$B\big(x, (x_n^*)\big) = \{\, y : y \in X,\ |x_n^*(y - x)| < 1 \text{ for each } n \,\}.$$

Prove the statements in (a) through (d).

(a) The collection of all sets $B\big(x, (x_n^*)\big)$ described in the above definition is a basis for a topology for X.

Let $\mathfrak{T}_0$ denote the topology from (a), and let $\mathfrak{T}_w$ and $\mathfrak{T}_{bw}$ denote the weak and bounded weak topologies of X respectively.

(b) The relative topologies inherited by each bounded subset of X from $\mathfrak{T}_w$ and $\mathfrak{T}_0$ are the same.

(c) $\mathfrak{T}_w \subseteq \mathfrak{T}_0 \subseteq \mathfrak{T}_{bw}$.

(d) The topology $\mathfrak{T}_0$ is completely regular and locally convex.

Instead of just trying to modify arguments from the proof of Theorem 2.7.2, it might be useful to consider $Q(X)$, where Q is the natural map from X into X^{**}.

2.88 The purpose of this exercise is to show that the topologies $\mathfrak{T}_{bw}$ and $\mathfrak{T}_0$ defined in Exercises 2.86 and 2.87 do not have to be the same. For each pair of positive integers m and n, let $x(m, n)$ be the element of c_0 that has its first $n-1$ terms equal to m^{-1}, its n^{th} term equal to m, and all the rest of its terms equal to 0. Let $A = \{ x(m,n) : m,n \in \mathbb{N} \}$.

(a) Show that A is b-weakly closed.

(b) Show that $0 \in \overline{A}^{\mathfrak{T}_0}$. Conclude that A is not $\mathfrak{T}_0$-closed, and therefore that $\mathfrak{T}_0 \subsetneqq \mathfrak{T}_{bw}$ for c_0.

This example is due to R. F. Wheeler [243].

2.8 Weak Compactness

Though Section 2.5 is devoted to the introduction and study of the weak topology of a normed space, nothing is said in that section about weakly compact sets other than that they are bounded. Actually, there is little more about weak compactness that could be said very easily in that section, since most of the interesting results about weakly compact sets are most readily obtained from facts about the weak* topology that are derived in Sections 2.6 and 2.7. Those results are now available, so it is time to take a closer look at weak compactness.

The first result of this section provides two tests for weak compactness using the weak* topology of the second dual.

2.8.1 Proposition. *Let A be a subset of a normed space X and let Q be the natural map from X into X^{**}. Then the following are equivalent.*

(a) *The set A is weakly compact.*

(b) *The set $Q(A)$ is weakly* compact.*

(c) *The set A is bounded and $Q(A)$ is weakly* closed.*

PROOF. By Proposition 2.6.24, the map Q is a weak-to-relative-weak* homeomorphism from X onto $Q(X)$. It follows that A is weakly compact if and only if $Q(A)$ is compact with respect to the relative weak* topology of $Q(X)$, which happens if and only if $Q(A)$ is weakly* compact in X^{**}. This proves the equivalence of (a) and (b). The equivalence of (b) and (c) is an easy consequence of the Banach-Alaoglu theorem, the boundedness of weakly* compact subsets of the dual space of a Banach space, and the fact that Q is an isometry. ■

The preceding result is one of the three key ingredients in the proof of the following important characterization of reflexivity, with the Banach-Alaoglu theorem and Goldstine's theorem being the other two.

2.8.2 Theorem. *A normed space is reflexive if and only if its closed unit ball is weakly compact.*

PROOF. Let Q be the natural map from a normed space X into X^{**}. It follows easily from Proposition 2.8.1, the Banach-Alaoglu theorem, and

Goldstine's theorem that each of statements 1 through 3 below is equivalent to the one following it.

1. The normed space X is reflexive.
2. $Q(B_X) = B_{X^{**}}$.
3. The set $Q(B_X)$ is weakly* closed.
4. The set B_X is weakly compact.

The equivalence of statements 1 and 4 is the conclusion of the theorem. ■

One nice property of topologies induced by metrics is that for such topologies, the properties of compactness, countable compactness, limit point compactness, and sequential compactness are equivalent. By a rather remarkable theorem due to W. F. Eberlein and V. L. Šmulian, the same is true for the weak topology of every normed space, despite the fact that this topology is not metrizable when the underlying space is infinite-dimensional; see Proposition 2.5.14. It is fairly easy to show that in any topological space, compactness $\Rightarrow$ countable compactness $\Rightarrow$ limit point compactness $\Leftarrow$ sequential compactness. The hard part of the proof of the Eberlein-Šmulian theorem is showing that weak limit point compactness implies both weak sequential compactness and weak compactness. The following lemmas will be useful for obtaining these latter implications.

2.8.3 Lemma. *Every relatively weakly limit point compact subset of a normed space is bounded.*

PROOF. Suppose that A is a relatively weakly limit point compact subset of a normed space X and that $x^* \in X^*$. By Corollary 2.5.6, it is enough to show that $x^*(A)$ is bounded.

Suppose to the contrary that $x^*(A)$ is not bounded. Let (x_n) be a sequence in A such that $|x^*x_{n+1}| \geq |x^*x_n| + 1$ for each n, and let w be a weak limit point of the infinite subset $\{x_n : n \in \mathbb{N}\}$ of A. Then the weak neighborhood $\{x : x \in X, |x^*x - x^*w| < \frac{1}{2}\}$ of w must contain two different members x_{n_0} and x_{n_1} of the sequence, which implies that $|x^*x_{n_0} - x^*x_{n_1}| < 1$. This contradiction proves the lemma. ■

2.8.4 Lemma. *Let X be a normed space and Y a finite-dimensional subspace of X^*. If $M > 1$, then there is a finite subset F_M of B_X such that $\|y^*\| \leq M \max\{|y^*x| : x \in F_M\}$ for each y^* in Y.*

PROOF. It may be assumed that $Y \neq \{0\}$. The compactness of S_Y implies that there is a finite subset $\{y_1^*, \dots, y_n^*\}$ of S_Y such that the open balls of radius $\frac{M-1}{2M}$ centered at $y_1^*, \dots, y_n^*$ cover S_Y. Since $\frac{M+1}{2M} < 1$, there are elements $x_1, \dots, x_n$ of B_X such that $|y_j^*x_j| > \frac{M+1}{2M}$ when $j = 1, \dots, n$. Let $F_M = \{x_1, \dots, x_n\}$.

Suppose that $y_0^* \in S_Y$. Select j so that $\|y_0^* - y_j^*\| < \frac{M-1}{2M}$. Then

$$\begin{aligned}|y_0^* x_j| &\geq |y_j^* x_j| - |y_0^* x_j - y_j^* x_j| \\ &\geq |y_j^* x_j| - \|y_0^* - y_j^*\| \, \|x_j\| \\ &> \tfrac{M+1}{2M} - \tfrac{M-1}{2M} \\ &= \tfrac{1}{M}.\end{aligned}$$

Therefore $\max\{\, |y_0^* x| : x \in F_M \,\} \geq \frac{1}{M}$. It follows that

$$M \max\{\, \|y^*\|^{-1} |y^* x| : x \in F_M \,\} \geq 1$$

whenever $y^* \in Y$ and $y^* \neq 0$, and therefore that

$$M \max\{\, |y^* x| : x \in F_M \,\} \geq \|y^*\|$$

for every y^* in Y. ∎

The following result is central to the proof of the Eberlein-Šmulian theorem to be given below. It is called Day's lemma because the argument used to prove it is essentially the same as an argument used by Mahlon Day to prove that every weakly sequentially compact subset of a normed space is weakly closed; see [54, Theorem III.2.4].

2.8.5 Day's Lemma. *Let X be a normed space.*

(a) *If A is a relatively weakly limit point compact subset of X and $x_0 \in \overline{A}^w$, then there is a sequence in A that converges weakly to x_0.*
(b) *If A_* is a relatively weakly limit point compact subset of X^* and $x_0^* \in \overline{A_*}^{w^*}$, then there is a sequence in A_* that converges weakly to x_0^*.*

Proof. The plan of attack is to prove (b) and then to obtain (a) from (b). To this end, let A_* be a relatively weakly limit point compact subset of X^* and let x_0^* be in $\overline{A_*}^{w^*}$. It may be assumed that $x_0^* \notin A_*$. By a straightforward argument involving various parts of Theorem 2.2.9 and the fact that $-x_0^* + A_*$ is relatively weakly limit point compact, it may also be assumed that $x_0^* = 0$.

The first order of business is to construct an increasing sequence (F_n) of nonempty finite subsets of B_X and a sequence (x_n^*) in A_* such that $x_{n_1}^* \neq x_{n_2}^*$ when $n_1 \neq n_2$ and such that for each positive integer n,

(1) $\|x^*\| \leq 2 \max\{\, |x^* x| : x \in F_n \,\}$ whenever $x^* \in \langle\, \{x_1^*, \dots, x_n^*\} \,\rangle$; and
(2) $\max\{\, |x_{n+1}^* x| : x \in F_n \,\} < \frac{1}{n+1}$.

To start the construction, let x_1^* be any element of A_*, let x_1 be a member of B_X such that $2|x_1^* x_1| \geq \|x_1^*\|$, and let $F_1 = \{x_1\}$. Now suppose that a nonempty finite subset F_n of B_X and elements $x_1^*, \ldots, x_n^*$ of A_* satisfy (1). Since $\{x^* : x^* \in X^*, |x^*x| < (n+1)^{-1}$ whenever $x \in F_n\}$ is a weak* neighborhood of 0 in X^* and $0 \in \overline{A_*}^{w^*} \setminus A_*$, there is an x_{n+1}^* in $A_* \setminus \{x_1^*, \ldots, x_n^*\}$ that satisfies (2). By Lemma 2.8.4, there is a finite subset F_n' of B_X such that $\|x^*\| \leq 2\max\{|x^*x| : x \in F_n'\}$ whenever $x^* \in \langle\{x_1^*, \ldots, x_{n+1}^*\}\rangle$. Letting $F_{n+1} = F_n \cup F_n'$ completes the inductive construction.

Let $D = \bigcup\{F_n : n \in \mathbb{N}\}$, a subset of B_X. Then (1) implies that $\|x^*\| \leq 2\sup\{|x^*x| : x \in D\}$ whenever $x^* \in \langle\{x_n^* : n \in \mathbb{N}\}\rangle$, from which it easily follows that $\|x^*\| \leq 2\sup\{|x^*x| : x \in D\}$ whenever $x^* \in [\{x_n^* : n \in \mathbb{N}\}]$. By hypothesis, the infinite subset $\{x_n^* : n \in \mathbb{N}\}$ of A_* has a weak limit point w^*. Since w^* must lie in the weakly closed set $[\{x_n^* : n \in \mathbb{N}\}]$, it follows that $\|w^*\| \leq 2\sup\{|w^*x| : x \in D\}$. For each x in D and each positive ϵ, the weak neighborhood $\{x^* : x^* \in X^*, |x^*x - w^*x| < \epsilon/2\}$ of w^* contains infinitely many members of $\{x_n^* : n \in \mathbb{N}\}$, which by (2) implies that $|w^*x| < \epsilon$ whenever $x \in D$ and $\epsilon > 0$, that is, that $w^*x = 0$ whenever $x \in D$. It follows that $w^* = 0$.

Thus, the set $\{x_n^* : n \in \mathbb{N}\}$ has 0 as its one and only weak limit point. If (x_n^*) did not converge weakly to 0, then there would be a weak neighborhood U of 0 and a subsequence $(x_{n_j}^*)$ of (x_n^*) lying entirely outside U. Then $\{x_{n_j}^* : j \in \mathbb{N}\}$ would have a weak limit point of its own, necessarily different from 0, even though this new weak limit point would also be a weak limit point of $\{x_n^* : n \in \mathbb{N}\}$ and therefore would have to be 0. This contradiction shows that $x_n^* \xrightarrow{w} 0$, which proves (b).

Now suppose that A is a relatively weakly limit point compact subset of X and that $x_0 \in \overline{A}^w$. Let Q be the natural map from X into X^{**}. Because of the weak-to-weak continuity of Q, the set $Q(A)$ is relatively weakly limit point compact in X^{**}. Since x_0 is the weak limit of a net from A, it follows that Qx_0 is the weak* limit of a net from $Q(A)$, so by (b) there is a sequence (x_n) in A such that $Qx_n \xrightarrow{w} Qx_0$. This implies that $x_n \xrightarrow{w} x_0$ and finishes the proof of (a). ■

As has already been mentioned, the hard parts of the proof of the following theorem are the arguments that (c) ⇒ (d) and that (c) ⇒ (a). These were, essentially, the respective contributions of Šmulian and Eberlein to this result. See [67, p. 466] for more details.

2.8.6 The Eberlein-Šmulian Theorem. (W. F. Eberlein, 1947 [71]; V. L. Šmulian, 1940 [225]). *Let A be a subset of a normed space. Then the following are equivalent.*

(a) *The set A is weakly compact.*

(b) *The set A is weakly countably compact.*

(c) *The set A is weakly limit point compact.*

(d) *The set A is weakly sequentially compact.*

Also, the following are equivalent.

(a_r) *The set A is relatively weakly compact.*

(b_r) *The set A is relatively weakly countably compact.*

(c_r) *The set A is relatively weakly limit point compact.*

(d_r) *The set A is relatively weakly sequentially compact.*

PROOF. It should be noted that the equivalence of (a_r) through (d_r) does not follow immediately from the equivalence of (a) through (d) by considering weak closures. The problem lies in the definition of relative limit point compactness and relative sequential compactness. It is possible for a subset of a topological space to be both relatively limit point compact and relatively sequentially compact, yet have its closure be neither limit point compact nor sequentially compact; see Exercise 2.15.

Let X be the normed space of which A is a subset. It is clear that (a) $\Rightarrow$ (b) and that (a_r) $\Rightarrow$ (b_r). Suppose that (c) does not hold, that is, that A has an infinite subset with no weak limit point in A. Then there is a countably infinite subset $\{ x_n : n \in \mathbb{N} \}$ of A with no weak limit point in A. For each positive integer n, let $U_n = X \setminus \overline{\{ x_j : j \geq n \}}^w$; notice that $A \setminus U_n = \{ x_j : j \geq n \}$. It follows that $\{ U_n : n \in \mathbb{N} \}$ is a countable covering of A by weakly open sets that cannot be thinned to a finite subcovering. The set A is therefore not weakly countably compact, which shows that (b) $\Rightarrow$ (c). Now suppose that the preceding argument is repeated under the stronger supposition that A has an infinite subset with no weak limit point in X. The set $\{ x_n : n \in \mathbb{N} \}$ will also have no weak limit point in X, which implies that $\{ x_j : j \geq n \}$ is weakly closed for each positive integer n. It follows that $U_n = X \setminus \{ x_j : j \geq n \}$ and $\overline{A}^w \setminus U_n = \{ x_j : j \geq n \}$ for each n. Therefore $\{ U_n : n \in \mathbb{N} \}$ is a countable covering of $\overline{A}^w$ by weakly open sets that cannot be thinned to a finite subcovering, so A is not relatively weakly countably compact. This proves that (b_r) $\Rightarrow$ (c_r). It is easy to see that (d) $\Rightarrow$ (c) and that (d_r) $\Rightarrow$ (c_r).

Notice that no special properties of the weak topology of X have yet been used, and in fact the preceding arguments show that (relative) compactness $\Rightarrow$ (relative) countable compactness $\Rightarrow$ (relative) limit point compactness $\Leftarrow$ (relative) sequential compactness in every topological space. What remains to be shown is that (relative) weak limit point compactness implies both (relative) weak sequential compactness and (relative) weak compactness in X.

For the rest of this proof, assume that A is relatively weakly limit point compact. The first order of business is to show that A is relatively weakly sequentially compact. Let (y_n) be a sequence in A. Since the goal is to show that (y_n) has a weakly convergent subsequence, it may be assumed that $y_{n_1} \neq y_{n_2}$ when $n_1 \neq n_2$. Let y_0 be a weak limit point of the infinite subset $\{ y_n : n \in \mathbb{N} \}$ of A. By discarding one y_n if necessary, it may be assumed that $y_0 \notin \{ y_n : n \in \mathbb{N} \}$. Since $y_0 \in \overline{\{ y_n : n \in \mathbb{N} \}}^w$, part (a) of Day's lemma assures that there is a sequence in $\{y_n : n \in \mathbb{N}\}$ that converges weakly to y_0. It follows easily that some subsequence of (y_n) converges weakly to y_0, which shows that $(c_r) \Rightarrow (d_r)$. If A were actually weakly limit point compact, then the weak limit point y_0 could have been selected to lie in A, which would have implied that (y_n) has a subsequence converging weakly to an element of A. It follows that (c) $\Rightarrow$ (d).

Let Q be the natural map from X into X^{**}. Since Q is weak-to-weak continuous and A is relatively weakly limit point compact, the set $Q(A)$ is also relatively weakly limit point compact. It follows from Lemma 2.8.3 that $Q(A)$ is bounded, and therefore that $\overline{Q(A)}^{w^*}$ is weakly* compact.

Suppose that $x_0^{**} \in \overline{Q(A)}^{w^*}$. By part (b) of Day's lemma, there is a sequence (z_n) in A such that $Qz_n \xrightarrow{w} x_0^{**}$. However, the relative weak sequential compactness of A assures that (z_n) has a subsequence converging weakly to a member z_0 of X. This implies that $x_0^{**} = Qz_0 \in Q(X)$, so $\overline{Q(A)}^{w^*} \subseteq Q(X)$. Since $Q^{-1}\left(\overline{Q(A)}^{w^*}\right)$ is a weakly compact subset of X and $A \subseteq Q^{-1}\left(\overline{Q(A)}^{w^*}\right)$, the set A is relatively weakly compact. This proves that $(c_r) \Rightarrow (a_r)$.

Finally, suppose that A is actually weakly limit point compact, and therefore weakly sequentially compact. If $w_0 \in \overline{A}^w$, then part (a) of Day's lemma produces a sequence in A that converges weakly to w_0. Since this sequence must in turn have a subsequence converging weakly to a point of A, it follows that $w_0 \in A$. Thus, the set A is weakly closed and relatively weakly compact, and so is weakly compact. This shows that (c) $\Rightarrow$ (a), the last implication needed to finish the proof of the theorem. ■

2.8.7 Corollary. *If A is a relatively weakly compact subset of a normed space and $x_0 \in \overline{A}^w$, then there is a sequence in A that converges weakly to x_0.*

PROOF. All that is needed is the observation that A is relatively weakly limit point compact, followed by an appeal to part (a) of Day's lemma. ■

The following corollary is often summarized by saying that weak compactness and relative weak compactness are separably determined. Notice that the condition in the statement of the corollary that $A \cap Y$ be weakly compact in Y is equivalent to $A \cap Y$ having that same property in X, and that the same is true for relative weak compactness since every closed subspace Y of X is a weakly closed subset of X.

2.8.8 Corollary. *Let A be a subset of a normed space X. Then A is weakly compact if and only if $A \cap Y$ is weakly compact in Y whenever Y is a separable closed subspace of X. The same is true if "weakly compact" is replaced by "relatively weakly compact."*

PROOF. If A is (relatively) weakly compact and Y is *any* closed subspace of X, then the intersection of A with the weakly closed set Y is (relatively) weakly compact in X, and hence with respect to the weak topology of Y.

Conversely, suppose that whenever Y is a separable closed subspace of X, the set $A \cap Y$ is (relatively) weakly compact in Y and therefore (relatively) weakly sequentially compact in Y. Let (x_n) be a sequence in A. Then $[\{x_n : n \in \mathbb{N}\}]$ is a closed subspace of X and is separable by Proposition 1.12.1 (a). It follows that (x_n) has a subsequence that is convergent with respect to the weak topologies of both Y and X. Notice that A must contain the weak limit if $A \cap [\{x_n : n \in \mathbb{N}\}]$ is actually weakly compact in $[\{x_n : n \in \mathbb{N}\}]$. It follows that A is (relatively) weakly sequentially compact and therefore (relatively) weakly compact. ■

The next two results were previously obtained in optional Section 1.13 from a sequential characterization of reflexivity by R. C. James. See Theorems 1.13.4, 1.13.5, and 1.13.8.

2.8.9 Corollary. *A normed space is reflexive if and only if each bounded sequence in the space has a weakly convergent subsequence.*

PROOF. A normed space is reflexive if and only if its closed unit ball is weakly compact, which by the Eberlein-Šmulian theorem and the fact that the closed unit ball is weakly closed is equivalent to the condition that every sequence in the closed unit ball has a weakly convergent subsequence. The corollary follows easily. ■

2.8.10 Corollary. *A normed space is reflexive if and only if each of its separable closed subspaces is reflexive.*

PROOF. Let X be a normed space. Then X is reflexive if and only if B_X is weakly compact, which by Corollary 2.8.8 is true if and only if the closed unit ball $B_X \cap Y$ of each separable closed subspace Y of X is weakly compact in Y, which holds if and only if each separable closed subspace of X is reflexive. ■

2.8.11 Corollary. *Every reflexive normed space is weakly sequentially complete.*

PROOF. Let (x_n) be a weakly Cauchy sequence in a reflexive normed space. Then (x_n) is bounded and so has a weakly convergent subsequence. It follows that (x_n) converges weakly to the weak limit of this subsequence. ■

2.8.12 Corollary. *A reflexive normed space has Schur's property if and only if it is finite-dimensional.*

PROOF. If X is a reflexive normed space with Schur's property, then every bounded sequence in X has a convergent subsequence, which implies that

X has the Heine-Borel property and therefore is finite-dimensional. Conversely, all finite-dimensional normed spaces have Schur's property since the weak and norm topologies of a finite-dimensional normed space are the same. ■

Suppose that K is a weakly compact subset of a Banach space X. Then the convex hull of K does not have to be weakly compact; see Exercise 2.93. As it turns out, all that really prevents $\mathrm{co}(K)$ from being weakly compact is that it might not be weakly closed, since it does have to be relatively weakly compact. (There can be further complications when X is not complete. See Exercise 2.94.)

The fact that the closed convex hull of a weakly compact subset of a Banach space must be weakly compact is due to M. G. Krein and V. L. Šmulian, and is the analog for the weak topology of an earlier result of S. Mazur for the norm topology. The two results have proofs that are very similar and that begin by considering the separable case.

2.8.13 Lemma. *Suppose that $\mathfrak{T}$ is the norm or weak topology of a separable Banach space X and that K is a $\mathfrak{T}$-compact subset of X. Then $\overline{\mathrm{co}}(K)$ is $\mathfrak{T}$-compact.*

PROOF. It may be assumed that $K \neq \emptyset$. Notice that K is weakly compact, whether $\mathfrak{T}$ is the norm or the weak topology. For the rest of this proof, the topology of K is assumed to be its relative $\mathfrak{T}$ topology whenever K is treated as a topological space.

This proof makes extensive use of the standard identification of $C(K)^*$ with $\mathrm{rca}(K)$; see Example 1.10.6. Let κ be the identity function on K, viewed as a map from the topological space K into the topological space consisting of X with the topology $\mathfrak{T}$. For each x^* in X^*, the map $x^*\kappa$ is in $C(K)$, so $\int_K x^*\kappa\, d\mu$ exists for each μ in $\mathrm{rca}(K)$.

Let μ_0 be a member of $\mathrm{rca}(K)$. It will be shown that there is a unique x_{μ_0} in X such that $x^*x_{\mu_0} = \int_K x^*\kappa\, d\mu_0$ for each x^* in X^*. To this end, suppose that (x_n^*) is a sequence in X^* that is weakly* convergent to some x_0^*. For each x in K and each positive integer n,

$$|x_n^*\kappa(x)| \leq \|x_n^*\|\,\|x\| \leq \sup\{\,\|x_m^*\| : m \in \mathbb{N}\,\}\,\sup\{\,\|y\| : y \in K\,\},$$

and Corollaries 2.6.10 and 2.5.7 imply that these last two suprema are finite. Since $x_n^*\kappa(x) \to x_0^*\kappa(x)$ for each x in K, it follows from Lebesgue's dominated convergence theorem that $\int_K x_n^*\kappa\, d\mu_0 \to \int_K x_0^*\kappa\, d\mu_0$. The linear functional $x^* \mapsto \int_K x^*\kappa\, d\mu_0$ on X^* is therefore weakly* sequentially continuous, and so is weakly* continuous by Corollary 2.7.10. Thus, there is a unique x_{μ_0} in X such that $x^*x_{\mu_0} = \int_K x^*\kappa\, d\mu_0$ for each x^* in X^*, as was claimed.

Define $T\colon C(K)^* \to X$ as follows. For each y^* in $C(K)^*$, let μ_{y^*} be the member of $\mathrm{rca}(K)$ identified with y^* and let $T(y^*)$ be the unique element $x_{\mu_{y^*}}$ of X such that $x^*x_{\mu_{y^*}} = \int_K x^*\kappa\, d\mu_{y^*}$ for each x^* in X^*. It is easy to check that T is linear. If (y^*_α) is a net in $C(K)^*$ that is weakly* convergent to some y^*, then for each x^* in X^*,

$$x^*Ty^*_\alpha = \int_K x^*\kappa\, d\mu_{y^*_\alpha} = y^*_\alpha(x^*\kappa) \to y^*(x^*\kappa) = x^*Ty^*,$$

and so $Ty^*_\alpha \xrightarrow{w} Ty^*$. Thus, the map T is weak*-to-weak continuous. It follows from this, the weak* compactness of $B_{C(K)^*}$, and the linearity of T that $T(B_{C(K)^*})$ is a weakly compact convex subset of X.

Suppose for the moment that $\mathfrak{T}$ is the norm topology. Let ϵ be a positive number. Since K is a compact metric space, there are elements $x_1, \dots, x_m$ of K and a partition $A_1, \dots, A_m$ of K into Borel subsets such that if each A_j has corresponding indicator function $\mathbf{I}_{A_j}$ and κ_ϵ is the function $\sum_{j=1}^m \mathbf{I}_{A_j} x_j$ from K into X, then $\|\kappa(x) - \kappa_\epsilon(x)\| < \epsilon$ for each x in K. Define $T_\epsilon \colon C(K)^* \to X$ as follows. For each y^* in $C(K)^*$, let μ_{y^*} be the member of $\mathrm{rca}(K)$ identified with y^* and let $T_\epsilon(y^*) = \sum_{j=1}^m \mu_{y^*}(A_j)x_j$. Then T_ϵ is linear and $\|T_\epsilon y^*\| \le \left(\sum_{j=1}^m \|x_j\|\right)\|y^*\|$ for each y^* in $C(K)^*$, so $T_\epsilon \in B\big(C(K)^*, X\big)$. If $x^* \in B_{X^*}$ and $y^* \in B_{C(K)^*}$, then

$$\begin{aligned}
|x^*(Ty^* - T_\epsilon y^*)| &= \left| \int_K x^*\kappa\, d\mu_{y^*} - \sum_{j=1}^m \mu_{y^*}(A_j)x^*x_j \right| \\
&= \left| \int_K x^*\kappa\, d\mu_{y^*} - \int_K x^*\kappa_\epsilon\, d\mu_{y^*} \right| \\
&= \left| \int_K x^*(\kappa - \kappa_\epsilon)\, d\mu_{y^*} \right| \\
&\le \|x^*\| \sup\{\, \|\kappa(x) - \kappa_\epsilon(x)\| : x \in K \,\}\, |\mu_{y^*}|(K) \\
&\le \epsilon.
\end{aligned}$$

Therefore, if $y^* \in B_{C(K)^*}$, then

$$\|Ty^* - T_\epsilon y^*\| = \max\{\, |x^*(Ty^* - T_\epsilon y^*)| : x^* \in B_{X^*} \,\} \le \epsilon.$$

Since T_ϵ is a bounded finite-rank linear operator, the subset $T_\epsilon(B_{C(K)^*})$ of X can be covered with a finite number of open balls of radius ϵ, and the inequality just proved assures that if the radius of each of these balls is doubled without changing the center, then the resulting open balls cover $T(B_{C(K)^*})$. The set $T(B_{C(K)^*})$ is therefore totally bounded, and so is norm compact since it is norm closed and the normed space X is complete.

Thus, the set $T(B_{C(K)^*})$ is a $\mathfrak{T}$-compact convex subset of X, whether $\mathfrak{T}$ is the norm or the weak topology.

Let x_0 be a member of K and let δ_{x_0} be the member of $\mathrm{rca}(K)$ such that $\delta_{x_0}(\{x_0\}) = 1$ and $|\delta_{x_0}|(K \setminus \{x_0\}) = 0$. Let y^*_0 be the member of $C(K)^*$ that corresponds to δ_{x_0}. Then for each x^* in X^*,

$$x^*Ty^*_0 = \int_K x^*\kappa\, d\delta_{x_0} = x^*\kappa(x_0) = x^*x_0,$$

so $Ty_0^* = x_0$. Since $\|y_0^*\| = \|\delta_{x_0}\| = 1$, it follows that $K \subseteq T(B_{C(K)^*})$, so $\overline{\text{co}}(K) \subseteq T(B_{C(K)^*})$. As a $\mathfrak{T}$-closed subset of a $\mathfrak{T}$-compact set, the set $\overline{\text{co}}(K)$ is itself $\mathfrak{T}$-compact. ∎

The following theorem was obtained by Krein under the additional assumption that X is separable. The general case is due to Krein and Šmulian.

2.8.14 The Krein-Šmulian Weak Compactness Theorem. (M. G. Krein, 1937 [143]; M. G. Krein and V. L. Šmulian, 1940 [146]). *The closed convex hull of a weakly compact subset of a Banach space is itself weakly compact.*

PROOF. Let K be a weakly compact subset of a Banach space X. If it can be shown that $\overline{\text{co}}(K)$ is weakly sequentially compact, then an application of the Eberlein-Šmulian theorem will finish the proof. Let (x_n) be a sequence in $\overline{\text{co}}(K)$. Since $\overline{\text{co}}(K)$ is weakly closed, it is enough to show that (x_n) has a weakly convergent subsequence. For this, notice that each x_n is the norm limit of a sequence of convex combinations of elements of K, from which it follows that there is a countable subset A of K such that

$$\{x_n : n \in \mathbb{N}\} \subseteq \overline{\text{co}}(A) \subseteq \overline{\text{co}}(K \cap [A]) \subseteq [A].$$

Since $K \cap [A]$ is a weakly compact subset of the separable Banach space $[A]$, it follows from the preceding lemma that $\overline{\text{co}}(K \cap [A])$ is weakly compact and thus weakly sequentially compact in $[A]$, and so (x_n) has a subsequence that is weakly convergent in $[A]$ and therefore in X. ∎

2.8.15 Mazur's Compactness Theorem. (S. Mazur, 1930 [161]). *The closed convex hull of a compact subset of a Banach space is itself compact.*

PROOF. Let K be a compact subset of a Banach space X. Then K is separable since it is a compact subset of a metric space, so $[K]$ is separable. It may therefore be assumed that X is itself separable. The conclusion of the theorem now follows immediately from Lemma 2.8.13. ∎

Recall that a normed space is *compactly generated* if it is the closed linear hull of one of its compact subsets, which is equivalent to the space being separable; see Theorem 1.12.15 and the comments following it. Since separable normed spaces have so many nice properties, it is natural to ask what special properties a normed space has when it is the closed linear hull of one of its *weakly* compact subsets, and in fact such spaces have been studied extensively. It turns out to be easier to prove theorems about spaces with this property when they are complete, which is why the completeness requirement is traditionally included in the following definition.

2.8.16 Definition. A normed space X is *weakly compactly generated* if it is a Banach space that includes a weakly compact subset K such that $X = [K]$.

Banach spaces that are the closed linear hull of a compact set or have a weakly compact closed unit ball are obviously weakly compactly generated, which immediately leads to two large classes of weakly compactly generated normed spaces.

2.8.17 Proposition. *If a Banach space is separable or reflexive, then it is weakly compactly generated.*

There are weakly compactly generated normed spaces that are neither separable nor reflexive; see Exercise 2.98. As is shown in Exercise 2.99, the space ℓ_∞ is not weakly compactly generated, so there are Banach spaces that lack this property.

See Mahlon Day's book [56, pp. 72–77] for more on weakly compactly generated normed spaces. Klaus Floret's book [79] is a good source of further information about weak compactness in general.

Exercises

2.89 Suppose that K is a weakly compact subset of the dual space of a normed space X.

(a) Prove that the relative topologies induced on K by the weak and weak* topologies of X^* are the same.

(b) Part (a) might seem to suggest that if $x^{**} \in X^{**}$, then there must be an x in X such that x^{**} and Qx agree on K, where Q is, as usual, the natural map from X into X^{**}. Show that this does not have to be so, even if K is norm compact and X is a Banach space.

2.90 Prove that a subset of ℓ_1 is compact if and only if it is weakly compact. Conclude that no infinite-dimensional subspace of ℓ_1 is reflexive.

2.91 Use the results of this section to prove that c_0 is not reflexive.

2.92 Let X be a normed space and let Q be the natural map from X into X^{**}.

(a) Find conditions on X necessary and sufficient for $Q(X)$ to be dense in X^{**}.

(b) Find conditions on X necessary and sufficient for $Q(X)$ to be weakly dense in X^{**}.

2.93 Find a compact subset K of a Banach space such that $\text{co}(K)$ is not closed. (A peek at Exercise 2.94 would not hurt.)

2.94 The goal of this exercise is to show that the conclusions of Lemma 2.8.13, the Krein-Šmulian weak compactness theorem, and Mazur's compactness theorem can all fail if the normed space in question is not required to be complete. Let X be the separable incomplete normed space consisting of the vector space of finitely nonzero sequences with the ℓ_∞ norm, and let $\{ e_n : n \in \mathbb{N} \}$ be the collection of standard unit vectors of this space. Show that the subset $\{0\} \cup \{ n^{-1}e_n : n \in \mathbb{N} \}$ of this space is compact, but that its closed convex hull is not even weakly compact.

2.95 (a) Prove that if a normed space X is weakly compactly generated, then there is a subset K of X that is convex, balanced, and weakly compact such that $X = \overline{\bigcup\{\, tK : t > 0\,\}}$.

(b) Suppose that the convex, balanced, weakly compact set K in (a) were such that $X = \bigcup\{\, tK : t > 0\,\}$. What could then be said about X?

2.96 (a) Prove that if a Banach space X is separable, then there is a subset K of X that is convex, balanced, and compact such that $X = \overline{\bigcup\{\, tK : t > 0\,\}}$.

(b) Suppose that the convex, balanced, compact set K in (a) were such that $X = \bigcup\{\, tK : t > 0\,\}$. What could then be said about X?

2.97 Prove that if $X_1, \dots, X_n$ are weakly compactly generated normed spaces, then $X_1 \oplus \cdots \oplus X_n$ is weakly compactly generated. Exercise 2.95 might help.

2.98 Give an example of a weakly compactly generated normed space that is neither separable nor reflexive. Exercises 1.145 and 2.97 might help.

2.99 (a) Prove that if K is a weakly compact subset of ℓ_∞, then the relative weak topology of K is induced by a metric. (Exercise 2.89 might help.) Conclude that K is norm separable.

(b) Prove that ℓ_∞ is not weakly compactly generated.

*2.9 James's Weak Compactness Theorem

This section is motivated by the form of James's theorem proved in Section 1.13, and requires the following material from that section: Lemmas 1.13.10, 1.13.12, and 1.13.13, and the observations about $L(x_n^*)$ made just after the proof of Theorem 1.13.11. Reading that material does not require the reading of any other part of that section. No other material from that or any other optional section is used in this section. In particular, James's theorem itself is not needed to prove any of the results of this section, and in fact will be derived from the main theorem of this section.

James's theorem says if X is a Banach space such that the supremum of $|x^*|$ on B_X is actually an attained maximum whenever $x^* \in X^*$, then X is reflexive. Robert James's first proofs of this for the separable and general cases appeared in 1957 [111] and 1964 [112] respectively, with a greatly simplified proof of the general case following in 1972 [117]. It is the 1972 proof that is presented in Section 1.13 of this book.

Since a Banach space is reflexive if and only if its closed unit ball is weakly compact, James's theorem can be interpreted to say that if the closed unit ball of a Banach space X has the property that $|x^*|$ attains its supremum on B_X whenever $x^* \in X^*$, then B_X is weakly compact. Motivated by this interpretation, Victor Klee conjectured in a 1962 paper [139] that James's result extends to every weakly closed subset of a Banach space; that is, that a weakly closed subset A of a Banach space X must be weakly compact if $\sup\{\,|x^*x| : x \in A\,\}$ is attained whenever $x^* \in X^*$. (A moment's thought about the half-open interval $(0, 1]$ in the Banach space $\mathbb{R}$ shows why A must be required to be closed in some sense. Also, requiring A to be closed in the norm sense is not enough; see Exercise 2.101.)

James proved Klee's conjecture in a 1964 paper [115] using some of the same ideas he had used in his proof of the reflexivity theorem. In his 1972 paper containing the simplified proof of the reflexivity theorem, James showed how that argument could be modified to obtain a simplified proof of the more general weak compactness result.

The purpose of this section is to give a proof of Klee's conjecture following James's lead from his 1972 paper. The general plan of attack is to use Lemma 1.13.10 to prove the weak compactness result for nonempty, separable, weakly closed subsets of the closed unit ball of a Banach space in Proposition 2.9.1, then use this proposition and Lemma 1.13.13 to obtain the corresponding result for nonempty, balanced, weakly closed subsets of the closed unit ball of a real Banach space in Proposition 2.9.2, and finally use Proposition 2.9.2 to obtain the general result and several useful variations of it in Theorem 2.9.3. This program parallels the one used in Section 1.13 to prove James's theorem, with Propositions 2.9.1 and 2.9.2 taking the place of Theorems 1.13.11 and 1.13.14 respectively. This is a full agenda, so it would be best to get started.

2.9.1 Proposition. (R. C. James, 1972 [117]). *Let A be a nonempty, separable, weakly closed subset of the closed unit ball of a Banach space X. Then the following are equivalent.*

(a) *The set A is not weakly compact.*

(b) *There is a θ for which $0 < \theta < 1$ and a sequence (x_n^*) in B_{X^*} such that $\lim_n x_n^*x = 0$ for each x in A and $\sup\{\,|x^*x| : x \in A\,\} \ge \theta$ whenever $x^* \in \operatorname{co}(\{\,x_n^* : n \in \mathbb{N}\,\})$.*

(c) *There is a θ for which $0 < \theta < 1$ such that if (β_n) is a sequence of positive numbers with sum 1, then there is an α such that $\theta \le \alpha \le 1$ and a sequence (y_n^*) in B_{X^*} such that*

  (1) $\lim_n y_n^*x = 0$ *for each x in A;*

  (2) $\sup\{\,|\sum_{j=1}^\infty \beta_j y_j^*x| : x \in A\,\} = \alpha$; *and*

  (3) $\sup\{\,|\sum_{j=1}^n \beta_j y_j^*x| : x \in A\,\} < \alpha(1 - \theta\sum_{j=n+1}^\infty \beta_j)$ *for each positive integer n.*

(d) *There is a z^* in X^* such that $\sup\{\,|z^*x| : x \in A\,\}$ is not attained.*

Proof. To see that (a) $\Rightarrow$ (b), suppose that A is not weakly compact. Let $V = [A]$ and let W be the vector space underlying V^* but with the norm given by the formula $\|v^*\|_W = \sup\{\,|v^*x| : x \in A\,\}$; it is easy to check that $\|\cdot\|_W$ really is a norm on W, with the key observation being that a member of W that is zero on A must be zero on V. Let $f\colon A \to W^*$ be defined by the formula $\big(f(x)\big)(v^*) = v^*x$; that is, let f be the "natural map" from A into W^*. Notice for this that the definition of $\|\cdot\|_W$ assures that $f(x)$ really is a bounded linear functional on W whenever $x \in A$, and that $f(A) \subseteq B_{W^*}$. Since V^* is a separating family of linear functionals on V, the function f is one-to-one. It is clear that a net (x_α) in A is weakly convergent in X to an x in A if and only if $f(x_\alpha) \xrightarrow{w^*} f(x)$ in W^*, from which it follows that f is a homeomorphism from A with its relative weak topology as a subset of X onto $f(A)$ with its relative weak* topology as a subset of W^*. Since A is not weakly compact in X, the set $f(A)$ is not weakly* compact in W^*. As a subset of the weakly* compact subset B_{W^*} of W^*, the set $f(A)$ must not be weakly* closed in W^*. Fix an element F of $\overline{f(A)}^{w^*} \setminus f(A)$. Notice that there cannot be a v in V such that $Fv^* = v^*v$ for each v^* in W, for it is easy to see that such a v would have to be in $\overline{A}^w \setminus A$, contradicting the fact that A is weakly closed. In particular, it follows that $F \neq 0$, and therefore that $\|F\|_{W^*} > 0$. Since $A \subseteq B_V$,

$$\begin{aligned}\sup\{\,|Fv^*| : v^* \in B_{V^*}\} &\le \sup\{\,|Fv^*| : v^* \in W,\ \sup\{\,|v^*x| : x \in A\,\} \le 1\,\}\\ &= \|F\|_{W^*},\end{aligned}$$

from which it follows that $F \in V^{**}$ and $\|F\|_{V^{**}} \le \|F\|_{W^*}$. Let Q_V be the natural map from V into V^{**}. Since V is complete, the set $Q_V(V)$ is closed in V^{**}, and since $F \notin Q_V(V)$, it follows that $d\big(F, Q_V(V)\big) > 0$, where d is determined by the norm of V^{**}. Let Δ be such that

$$0 < \Delta < d\big(F, Q_V(V)\big)$$

and let $\{\,a_n : n \in \mathbb{N}\,\}$ be a countable dense subset of A. If $n \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_{n+1}$ are scalars, then

$$\Big|\alpha_1\Delta + \sum_{j=1}^{n} \alpha_{j+1}0\Big| = |\alpha_1|\Delta \le \frac{\Delta}{d\big(F, Q_V(V)\big)}\Big\|\alpha_1F + \sum_{j=1}^{n} \alpha_{j+1}Q_Va_j\Big\|_{V^{**}},$$

which by Helly's theorem implies the existence of a v_n^* in V^* such that

(i) $\displaystyle \|v_n^*\|_{V^*} \le \frac{\Delta}{d\big(F, Q_V(V)\big)} + \frac{d\big(F, Q_V(V)\big) - \Delta}{2\cdot d\big(F, Q_V(V)\big)} < 1;$

(ii) $Fv_n^* = \Delta$; and

(iii) $v_n^*a_j = (Q_Va_j)v_n^* = 0$ if $j \le n$.

For each n, let x_n^* be a Hahn-Banach extension of v_n^* to X. Then (iii) together with the density of $\{\,a_n : n \in \mathbb{N}\,\}$ in A assures that $\lim_n x_n^*x = 0$ for each x in A. If $x^* \in \mathrm{co}(\{\,x_n^* : n \in \mathbb{N}\,\})$ and v^* is the restriction of x^* to V, then

$$\Delta = Fv^* \le \|F\|_{W^*}\sup\{\,|v^*x| : x \in A\,\} = \|F\|_{W^*}\sup\{\,|x^*x| : x \in A\,\}.$$

Let $\theta = \Delta/\|F\|_{W^*}$. Since

$$0 < \Delta = Fv_1^* \leq \|F\|_{V^{**}}\|v_1^*\|_{V^*} < \|F\|_{W^*},$$

it follows that $0 < \theta < 1$. Therefore θ and (x_n^*) satisfy (b), which finishes the proof that (a) $\Rightarrow$ (b).

Now assume the existence of a θ and a sequence (x_n^*) as in (b), and let (β_n) be a sequence of positive reals summing to 1. Then the α and the sequence (y_n^*) guaranteed by Lemma 1.13.10 do all that is required of them in (c). In particular, notice that if $x \in A$, then the facts that $y_n^* \in \operatorname{co}(\{\, x_j^* : j \geq n \,\})$ for each n and $\lim_n x_n^* x = 0$ together imply that $\lim_n y_n^* x = 0$. It follows that (b) $\Rightarrow$ (c).

Suppose that (c) holds. Fix a sequence (β_n) of positive scalars summing to 1. Let α and (y_n^*) be as in (c) and let $z^* = \sum_{j=1}^\infty \beta_j y_j^*$. It will be shown that $\sup\{\, |z^*x| : x \in A \,\}$ is not attained. Let x_0 be an element of A and let n be a positive integer such that $|y_j^* x_0| < \alpha\theta$ whenever $j > n$. Then

$$\begin{aligned}
|z^*x_0| &= \left|\sum_{j=1}^\infty \beta_j y_j^* x_0\right| \\
&\leq \left|\sum_{j=1}^n \beta_j y_j^* x_0\right| + \sum_{j=n+1}^\infty \beta_j |y_j^* x_0| \\
&< \sup\left\{ \left|\sum_{j=1}^n \beta_j y_j^* x\right| : x \in A \right\} + \alpha\theta \sum_{j=n+1}^\infty \beta_j \\
&< \alpha\left(1 - \theta \sum_{j=n+1}^\infty \beta_j\right) + \alpha\theta \sum_{j=n+1}^\infty \beta_j \\
&= \alpha \\
&= \sup\{\, |z^*x| : x \in A \,\},
\end{aligned}$$

so $\sup\{\, |z^*x| : x \in A \,\}$ is not attained at x_0. This proves that (c) $\Rightarrow$ (d).

Finally, if A is weakly compact and $x^* \in X^*$, then the weakly continuous function $|x^*|$ must attain its supremum on A, so (d) $\Rightarrow$ (a). ■

For the next result, it will be necessary to reinstate temporarily some notation used in Section 1.13. Suppose that X is a *real* normed space and that (x_n^*) is a bounded sequence in X^*. Let

$$L(x_n^*) = \{\, x^* : x^* \in X^*,\ x^*x \leq \limsup_n x_n^* x \text{ whenever } x \in X \,\}.$$

The only use that will be made of this notation in the rest of this book is in the next proposition. It was shown in Lemma 1.13.12 that $L(x_n^*) \neq \varnothing$. Notice also that $\liminf_n x_n^* x \leq x^* x$ whenever $x^* \in L(x_n^*)$ and $x \in X$, since $\liminf_n x_n^* x = -\limsup_n x_n^*(-x)$.

2.9.2 Proposition. (R. C. James, 1972 [117]). *Let A be a nonempty, balanced, weakly closed subset of the closed unit ball of a real Banach space X. Then the following are equivalent.*

(a) *The set A is not weakly compact.*

(b) *There is a θ for which $0 < \theta < 1$, a subset A_0 of A, and a sequence (x_n^*) in B_{X^*} such that $\sup\{\,|(x^* - w^*)(x)| : x \in A\,\} \geq \theta$ for each x^* in $\operatorname{co}(\{\,x_n^* : n \in \mathbb{N}\,\})$ and w^* in $A_0^\perp$, and $\lim_n x_n^* x = 0$ for each x in A_0.*

(c) *There is a θ for which $0 < \theta < 1$ such that if (β_n) is a sequence of positive numbers with sum 1, then there is an α such that $\theta \leq \alpha \leq 2$ and a sequence (y_n^*) in B_{X^*} such that whenever $w^* \in L(y_n^*)$,*
   (1) $\sup\{\,|\sum_{j=1}^\infty \beta_j (y_j^* - w^*)(x)| : x \in A\,\} = \alpha$; *and*
   (2) $\sup\{\,|\sum_{j=1}^n \beta_j (y_j^* - w^*)(x)| : x \in A\,\} < \alpha(1 - \theta \sum_{j=n+1}^\infty \beta_j)$ *for each positive integer n.*

(d) *There is a z^* in X^* such that $\sup\{\,|z^*x| : x \in A\,\}$ is not attained.*

PROOF. Suppose first that A is not weakly compact. By Corollary 2.8.8, there is a separable closed subspace Y of X such that $A \cap Y$ is not weakly compact. It follows from Proposition 2.9.1 that there is a θ for which $0 < \theta < 1$ and a sequence (x_n^*) in B_{X^*} such that $\lim_n x_n^* x = 0$ for each x in $A \cap Y$ and $\sup\{\,|x^*x| : x \in A \cap Y\,\} \geq \theta$ whenever $x^* \in \operatorname{co}(\{\,x_n^* : n \in \mathbb{N}\,\})$. Let $A_0 = A \cap Y$. If $x^* \in \operatorname{co}(\{\,x_n^* : n \in \mathbb{N}\,\})$ and $w^* \in A_0^\perp$, then

$$\begin{aligned}\sup\{\,|(x^* - w^*)(x)| : x \in A\,\} &\geq \sup\{\,|(x^* - w^*)(x)| : x \in A_0\,\}\\ &= \sup\{\,|x^*x| : x \in A_0\,\} \geq \theta.\end{aligned}$$

The number θ, the set A_0, and the sequence (x_n^*) do all that is required of them in (b), so (a) $\Rightarrow$ (b).

Now suppose that (b) holds. Since $\liminf_n x_n^* x \leq x^* x \leq \limsup_n x_n^* x$ whenever $x^* \in L(x_n^*)$ and $x \in X$, it follows that $x^* x = \lim_n x_n^* x = 0$ whenever $x^* \in L(x_n^*)$ and $x \in A_0$, that is, that $L(x_n^*) \subseteq A_0^\perp$. This and Lemma 1.13.13 together imply that (c) holds, so (b) $\Rightarrow$ (c).

Suppose next that (c) holds. Let Δ be a scalar such that $0 < \Delta < \theta^2/2$. For each positive integer n, let

$$\beta_n = \frac{2 - \Delta}{\Delta} \left(\frac{\Delta}{2}\right)^n.$$

Then (β_n) is a sequence of positive scalars that sums to 1. Let α and (y_n^*) be as in (c). Let w^* be any member of $L(y_n^*)$ and let $z^* = \sum_{j=1}^\infty \beta_j (y_j^* - w^*)$. It will be shown that $\sup\{\,|z^*x| : x \in A\,\}$ is not attained. To this end, suppose that $x_0 \in A$. Since $\liminf_j y_j^* x_0 \leq w^* x_0$ and $\theta \leq \alpha$, there is an n such that

$$(y_{n+1}^* - w^*)(x_0) < \theta^2 - 2\Delta \leq \alpha\theta - 2\Delta.$$

Since $w^*x \le \limsup_j y_j^*x \le 1$ whenever $x \in B_X$, it follows that $\|w^*\| \le 1$. Therefore

$$\begin{aligned}
z^*x_0 &= \sum_{j=1}^{\infty} \beta_j(y_j^* - w^*)(x_0) \\
&< \sum_{j=1}^{n} \beta_j(y_j^* - w^*)(x_0) + (\alpha\theta - 2\Delta)\beta_{n+1} + \sum_{j=n+2}^{\infty} \beta_j(y_j^* - w^*)(x_0) \\
&\le \sup\left\{ \left|\sum_{j=1}^{n} \beta_j(y_j^* - w^*)(x)\right| : x \in A \right\} + (\alpha\theta - 2\Delta)\beta_{n+1} + 2\sum_{j=n+2}^{\infty} \beta_j \\
&< \alpha\left(1 - \theta\sum_{j=n+1}^{\infty} \beta_j\right) + (\alpha\theta - 2\Delta)\beta_{n+1} + 2\sum_{j=n+2}^{\infty} \beta_j.
\end{aligned}$$

Since $\sum_{j=n+2}^{\infty} \beta_j = \frac{1}{2}\Delta \sum_{j=n+1}^{\infty} \beta_j < \Delta \sum_{j=n+1}^{\infty} \beta_j$,

$$\begin{aligned}
z^*x_0 &< \alpha - (\alpha\theta - 2\Delta)\sum_{j=n+1}^{\infty} \beta_j + (\alpha\theta - 2\Delta)\beta_{n+1} \\
&= \alpha - (\alpha\theta - 2\Delta)\sum_{j=n+2}^{\infty} \beta_j \\
&< \alpha \\
&= \sup\left\{ \left|\sum_{j=1}^{\infty} \beta_j(y_j^* - w^*)(x)\right| : x \in A \right\} \\
&= \sup\{\, |z^*x| : x \in A \,\}.
\end{aligned}$$

Since A is balanced and therefore contains $-x_0$, it is also true that $-z^*x_0 < \sup\{\, |z^*x| : x \in A \,\}$, and therefore that $|z^*x_0| < \sup\{\, |z^*x| : x \in A \,\}$. This proves that $\sup\{\, |z^*x| : x \in A \,\}$ is not attained, and shows that (c) $\Rightarrow$ (d).

Finally, if $z^* \in X^*$ and the weakly continuous function $|z^*|$ does not attain its supremum on A, then A cannot be weakly compact, which proves that (d) $\Rightarrow$ (a). ■

2.9.3 James's Weak Compactness Theorem. (R. C. James, 1964 [115]). *Suppose that A is a nonempty weakly closed subset of a Banach space X. Then the following are equivalent.*

(a) *The set A is weakly compact.*

(b) *Whenever x^* is a bounded linear functional on X, the supremum of $|x^*|$ on A is attained.*

(c) *Whenever u^* is a bounded real-linear functional on X, the supremum of $|u^*|$ on A is attained.*

(d) *Whenever u^* is a bounded real-linear functional on X, the supremum of u^* on A is attained.*

PROOF. The weak continuity of the functions mentioned in (b), (c), and (d) assures that (a) implies each of (b), (c), and (d).

For the rest of this proof, it may be assumed that A is bounded, for the hypotheses of (b), (c), and (d) all assure that $x^*(A)$ is a bounded set of scalars for each x^* in X^*, and therefore, by Corollary 2.5.6, that A is bounded. Clearly, it may then be assumed that $A \subseteq B_X$. Let X_r be the Banach space formed from X by using real scalars (which does not imply any assumption that $\mathbb{F} = \mathbb{C}$; if $\mathbb{F} = \mathbb{R}$, then X_r and X are the same Banach space). Even when $\mathbb{F} = \mathbb{C}$, straightforward arguments based on Proposition 1.9.3 and Corollary 2.1.22 show that the weak topologies of X_r and X are the same topology for the set underlying these Banach spaces.

Suppose that (b) holds. Let

$$B = \bigcap_{x^* \in X^*} \{\, x : x \in X,\ |x^*x| \le \sup\{\, |x^*y| : y \in A \,\} \,\}.$$

Then B is a balanced, weakly closed subset of B_X that includes A, and $\sup\{\, |x^*x| : x \in B \,\}$ is attained by $|x^*|$ on B, in fact on A, whenever $x^* \in X^*$. If $\mathbb{F} = \mathbb{C}$, then easy arguments based on the fact that B is balanced and on standard facts about the relationship between real-linear and complex-linear functionals show that

$$\sup\{\, |\mathrm{Re}\, x^*x| : x \in B \,\} = \sup\{\, |x^*x| : x \in B \,\}$$

for each x^* in X^* and that $\sup\{\, |u^*x| : x \in B \,\}$ is attained whenever $u^* \in (X_r)^*$. Whether or not $\mathbb{F} = \mathbb{C}$, it now follows from Proposition 2.9.2 that B is a weakly compact subset of X_r, and so of X, and therefore that the weakly closed subset A of B is also weakly compact. This proves that (b) $\Rightarrow$ (a).

Suppose next that (c) holds. Since it has already been proved that (b) $\Rightarrow$ (a) when $\mathbb{F} = \mathbb{R}$, it follows immediately that A is a weakly compact subset of X_r and therefore of X, which shows that (c) $\Rightarrow$ (a).

Finally, suppose that (d) holds. If u^* is a bounded real-linear functional on X, then

$$\sup\{\, |u^*x| : x \in A \,\} = \max\bigl\{\sup\{\, u^*x : x \in A \,\}, \sup\{\, -u^*x : x \in A \,\}\bigr\},$$

from which it follows that $\sup\{\, |u^*x| : x \in A \,\}$ is attained. This shows that (d) $\Rightarrow$ (c) and finishes the proof of the theorem. ■

James's theorem is itself an easy corollary of James's weak compactness theorem.

2.9.4 James's Theorem. (R. C. James, 1964 [112]). *If every bounded linear functional on a Banach space is norm-attaining, then the space is reflexive.*

PROOF. Just observe that if every bounded linear functional on a Banach space is norm-attaining, then James's weak compactness theorem implies that the closed unit ball of the space must be weakly compact. ■

Of course, if a Banach space X is reflexive, then for each x^* in X^* the weakly continuous function $|x^*|$ attains its supremum on the weakly compact set B_X, so the condition that all bounded linear functionals be norm-attaining is actually equivalent to reflexivity for Banach spaces.

In 1971, James [116] gave an example of a normed space, necessarily incomplete, with the property that each bounded linear functional on the space is norm-attaining, though the space is not reflexive and its closed unit ball is therefore not weakly compact. Thus, the completeness requirement cannot be deleted from the hypotheses of either James's theorem or James's weak compactness theorem. See also Exercise 2.102.

Actually, the term *James's theorem* is often used for both the reflexivity theorem proved immediately above and the weak compactness theorem that also bears James's name. In practice, the context always makes it clear which of these two closely related results is intended.

Exercises

2.100 Either prove the following weak* analog of James's weak compactness theorem or find a counterexample: *Suppose that A is a nonempty weakly* closed subset of the dual space of a Banach space X. Then A is weakly* compact if and only if* $\sup\{\,|x^*x| : x^* \in A\,\}$ *is attained whenever* $x \in X$.

2.101 Find a nonempty closed subset A of a Banach space X such that A is not weakly compact even though the supremum of $|x^*|$ on A is attained whenever $x^* \in X^*$.

2.102 Without citing the 1971 example by James mentioned in this section, find a nonempty weakly closed subset A of an incomplete normed space X such that A is not weakly compact even though the supremum of $|x^*|$ on A is attained whenever $x^* \in X^*$. Exercise 2.94 might help.

2.103 Suppose that X is a Banach space having a subset A with nonempty interior such that the supremum of $|x^*|$ on A is attained whenever $x^* \in X^*$. Prove that X is reflexive.

2.104 Prove the following result.

Theorem. (V. L. Šmulian, 1939 [223]). *A convex subset C of a Banach space is weakly compact if and only if $\bigcap_n C_n \neq \varnothing$ whenever (C_n) is a decreasing*[6] *sequence of nonempty convex subsets of C that are closed in C.*

---

[6]Recall that a sequence (A_n) of subsets of a set X is *decreasing* if $A_n \supseteq A_{n+1}$ for each n, even if some or all of the inclusions are equalities.

2.105 Prove the following result.

Theorem. (V. L. Klee, 1962 [139]). *Suppose that C is a closed bounded non-weakly-compact convex subset of a Banach space X. Then there is a decreasing sequence (C_n) of nonempty closed convex subsets of C such that whenever $x \in C$ and $0 \leq t < 1$, the set $x + t(-x + C)$ intersects only finitely many of the sets C_n.*

2.106 Derive the Krein-Šmulian weak compactness theorem from James's weak compactness theorem.

2.10 Extreme Points

Suppose that C is a closed convex polygonal region in the Euclidean plane and that $v_1, \ldots, v_n$ are the vertices of the polygon that bounds C. Then it is easy to see that C is the convex hull of the set $\{v_1, \ldots, v_n\}$. The main purpose of this section is to show that the same is true whenever C is a nonempty compact convex subset of a Hausdorff locally convex space, provided that the notion of a "vertex" of the boundary of C is appropriately generalized and that the *closed* convex hull of the set of "vertices" is used. As one consequence, it will be seen that certain common Banach spaces are not isometrically isomorphic to the dual space of any normed space.

2.10.1 Definition. Let C be a nonempty closed convex subset of a TVS. A subset D of C is *extremal* in C if D is itself a nonempty closed convex set having this additional property: If $x, y \in C$ and $tx + (1 - t)y \in D$ for some t such that $0 < t < 1$, then $x, y \in D$.

That is, a nonempty closed convex subset D of C is extremal in C if every "closed line segment" in C whose "interior" intersects D must lie entirely in D.

2.10.2 Definition. Let C be a nonempty closed convex subset of a TVS that is Hausdorff (so that singleton sets are closed). An *extreme point* of C is an element x of C such that $\{x\}$ is an extremal subset of C.

In other words, an element of C is an extreme point of C if the point does not lie in the "interior" of any nontrivial "closed line segment" in C.

2.10.3 Example. In real ℓ_∞^2, the extreme points of the closed unit ball are the four points $(\pm 1, \pm 1)$, that is, the vertices of the polygonal boundary of the closed unit ball. In real Euclidean 2-space, the set of extreme points of the closed unit ball is the entire unit sphere.

2.10.4 Example. Suppose that (α_n) is an element of the closed unit ball of c_0. Let n_0 be a positive integer such that $|\alpha_{n_0}| < \frac{1}{2}$, and let (β_n) and (γ_n)

be the members of c_0 such that $\beta_n = \gamma_n = \alpha_n$ if $n \neq n_0$, but $\beta_{n_0} = \alpha_{n_0} + \frac{1}{2}$ and $\gamma_{n_0} = \alpha_{n_0} - \frac{1}{2}$. Then (β_n) and (γ_n) are different elements of B_{c_0} such that $(\alpha_n) = \frac{1}{2}(\beta_n) + \frac{1}{2}(\gamma_n)$, so (α_n) is not an extreme point of B_{c_0}. Therefore B_{c_0} has no extreme points, so not even in Banach spaces are bounded nonempty closed convex sets guaranteed to have extreme points.

Since c_0 is infinite-dimensional and not reflexive, its closed unit ball is neither norm compact nor weakly compact, but this leaves open the possibility that there might be some other Hausdorff locally convex topology for c_0 with respect to which B_{c_0} is compact. One of the consequences of the Krein-Milman theorem, the main result of this section, is that this possibility is ruled out, for the compactness of B_{c_0} with respect to a Hausdorff locally convex topology for c_0 would force B_{c_0} to have extreme points.

The following collection of simple facts about extremal sets will be useful in the proof of the Krein-Milman theorem.

2.10.5 Lemma. *Let C be a nonempty closed convex subset of a TVS X.*

(a) *If D is an extremal subset of C and D' is an extremal subset of D, then D' is an extremal subset of C.*

(b) *If $\mathfrak{F}$ is a nonempty family of extremal subsets of C and $\bigcap\{ D : D \in \mathfrak{F} \}$ is nonempty, then $\bigcap\{ D : D \in \mathfrak{F} \}$ is an extremal subset of C.*

(c) *Suppose that C is compact and that $x^* \in X^*$. Then*

$$\{ x : x \in C, \operatorname{Re} x^*x = \max\{ \operatorname{Re} x^*y : y \in C \} \}$$

is an extremal subset of C.

PROOF. Parts (a) and (b) follow easily from Definition 2.10.1. For (c), suppose that C is compact and that $x^* \in X^*$. Let $s = \max\{ \operatorname{Re} x^*x : x \in C \}$ and let $D_s = \{ x : x \in C, \operatorname{Re} x^*x = s \}$, a nonempty closed convex subset of C. If x and y are members of C such that $tx + (1-t)y \in D_s$ for some t in $(0,1)$, then $t \operatorname{Re} x^*x + (1-t) \operatorname{Re} x^*y = s$, which together with the definition of s implies that $\operatorname{Re} x^*x = \operatorname{Re} x^*y = s$ and thus that $x, y \in D_s$. The set D_s is therefore an extremal subset of C. ■

2.10.6 The Krein-Milman Theorem. (M. G. Krein and D. P. Milman, 1940 [145]). *Let C be a nonempty compact convex subset of a Hausdorff LCS. Then C is the closed convex hull of its set of extreme points.*

PROOF. Let X be the Hausdorff LCS of which C is a subset. The first order of business is to show that C has at least one extreme point. To this end, let $\mathfrak{D}$ be the collection of all extremal subsets of C, preordered by reverse inclusion; that is, preordered by declaring that $D_1 \preceq D_2$ when $D_1 \supseteq D_2$. Then $\mathfrak{D} \neq \emptyset$ since $C \in \mathfrak{D}$. If $\mathfrak{C}$ is a nonempty chain in $\mathfrak{D}$, then it follows from the compactness of C that $\bigcap\{ D : D \in \mathfrak{C} \} \neq \emptyset$ and therefore from

Lemma 2.10.5 (b) that $\bigcap\{ D : D \in \mathfrak{C} \}$ is an upper bound for $\mathfrak{C}$ in $\mathfrak{D}$. By Zorn's lemma, the set $\mathfrak{D}$ has a maximal element M.

Suppose that M were to contain two distinct elements y and z. By Corollary 2.2.22, there would be an x^* in X^* such that $x^*y \neq x^*z$; after multiplying x^* by i if necessary, it may be assumed that $\operatorname{Re} x^*y \neq \operatorname{Re} x^*z$. Since $\operatorname{Re} x^*$ is not constant on M, it follows from parts (a) and (c) of Lemma 2.10.5 that $\{ x : x \in M, \operatorname{Re} x^*x = \max\{ \operatorname{Re} x^*y : y \in M \} \}$ would be a proper subset of M that is also an extremal subset of C, which contradicts the maximality of M. Thus, the set M contains only one point, and so C has an extreme point.

Let E be the collection of all extreme points of C. To finish the proof, it is enough to show that $C \subseteq \overline{\operatorname{co}}(E)$. Suppose to the contrary that $C \not\subseteq \overline{\operatorname{co}}(E)$. Applying Theorem 2.2.28 to the compact set $\overline{\operatorname{co}}(E)$ and any one-element subset $\{x_0\}$ of $C \setminus \overline{\operatorname{co}}(E)$ yields an x_0^* in X^* such that

$$\max\{ \operatorname{Re} x_0^*x : x \in \overline{\operatorname{co}}(E) \} < \operatorname{Re} x_0^*x_0 \leq m_0 = \max\{ \operatorname{Re} x_0^*x : x \in C \}.$$

The extremal subset $\{ x : x \in C, \operatorname{Re} x_0^*x = m_0 \}$ of C is itself a nonempty compact convex subset of X, and therefore has an extreme point x_e which must also be an extreme point of C. However, the fact that $\operatorname{Re} x_0^*x_e = m_0 > \sup\{ \operatorname{Re} x_0^*x : x \in E \}$ prevents x_e from being an extreme point of C, a contradiction that finishes the proof of the theorem. ■

2.10.7 Corollary. *Let C be a nonempty compact convex subset of a Hausdorff LCS X and let x^* be a member of X^*. Then there is an extreme point x_e of C such that $\operatorname{Re} x^*x_e = \max\{ \operatorname{Re} x^*x : x \in C \}$. Consequently, if E is the set of extreme points of C, then $\max\{ \operatorname{Re} x^*x : x \in C \} = \max\{ \operatorname{Re} x^*x : x \in E \}$.*

PROOF. The extremal subset $\{ x : x \in C, \operatorname{Re} x^*x = \max\{ \operatorname{Re} x^*y : y \in C \} \}$ of C is a nonempty compact convex subset of X, and therefore has an extreme point x_e which is also an extreme point of C. The corollary follows immediately from this. ■

2.10.8 Corollary. *If X is a reflexive normed space, then B_X is the closed convex hull of its set of extreme points.*

PROOF. Since the closed convex hulls of the set of extreme points of B_X with respect to the norm and weak topologies are the same, it is enough to know that B_X is weakly compact, which follows from the reflexivity of X. ■

2.10.9 Corollary. *If X is a normed space, then B_{X^*} is the weakly* closed convex hull of its set of extreme points.*

PROOF. This follows immediately from the weak* compactness of B_{X^*}. ■

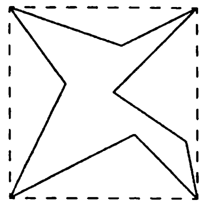


FIGURE 2.2. A polygon P in the Euclidean plane, and its closed convex hull.

2.10.10 Corollary. *An infinite-dimensional normed space whose closed unit ball has only finitely many extreme points is not isometrically isomorphic to the dual space of any normed space.*

PROOF. It is clear that isometric isomorphisms from one normed space onto another preserve extreme points, so it is enough to show that the closed unit ball of an infinite-dimensional normed space's dual space always has infinitely many extreme points. Let X be an infinite-dimensional normed space. Then X^* is also infinite-dimensional, and B_{X^*} is the weakly* closed convex hull of its set E of extreme points. The weakly* closed convex hull of a finite subset of X^* must lie in the finite-dimensional linear hull of that finite set, and therefore cannot be B_{X^*}. Thus, the set E is infinite. ∎

2.10.11 Example. As was shown in Example 2.10.4, the closed unit ball of c_0 has no extreme points, so c_0 is not isometrically isomorphic to the dual space of any normed space. Similarly, the space $L_1[0,1]$ is not isometrically isomorphic to the dual space of any normed space, since it can be shown that $B_{L_1[0,1]}$ has no extreme points. See Exercise 2.110.

Suppose that P is a polygon in the Euclidean plane, not assumed to be convex, whose sides intersect only at its vertices and whose vertices are each common to exactly two sides; see Figure 2.2. Let K be the compact region in the plane consisting of P together with the bounded component of its complement. It is easy to see that $\overline{\text{co}}(K)$ (which is the same as $\text{co}(K)$ in this case) is a closed convex polygonal region whose vertices are a subset of the vertices of K. In particular, the extreme points of $\overline{\text{co}}(K)$ all lie in K. There is a partial converse of the Krein-Milman theorem, due to Milman, that says that the same is true whenever K is a nonempty compact subset of a Hausdorff LCS, provided that $\overline{\text{co}}(K)$ is also compact. To get this result, several lemmas will be used that are of some interest in their own right.

2.10.12 Lemma. *Suppose that x_e is an extreme point of a nonempty closed convex subset C of a Hausdorff TVS and that $x_e = \sum_{j=1}^{n} t_j x_j$, where $t_1, \dots, t_n$ are nonnegative real numbers summing to 1 and $x_1, \dots, x_n \in C$. Then $x_j = x_e$ for each j such that $t_j \neq 0$.*

PROOF. It may be assumed that each t_j is nonzero and that $n \geq 2$. Fix an integer j_0 such that $1 \leq j_0 \leq n$. Then $\sum_{j \neq j_0} (1 - t_{j_0})^{-1} t_j x_j \in C$ because

C is convex. Since x_e is an extreme point of C and

$$x_e = t_{j_0}x_{j_0} + (1 - t_{j_0})\sum_{j\neq j_0}(1 - t_{j_0})^{-1}t_jx_j,$$

it follows that $x_{j_0} = x_e$. ■

2.10.13 Lemma. *Suppose that $C_1, \ldots, C_n$ are nonempty convex subsets of a vector space. Then $\mathrm{co}(C_1 \cup \cdots \cup C_n)$ consists of all sums $\sum_{j=1}^n t_jx_j$ such that $t_1, \ldots, t_n$ are nonnegative real numbers summing to 1 and $x_j \in C_j$ when $j = 1, \ldots, n$.*

PROOF. It suffices to show that every member of $\mathrm{co}(C_1 \cup \cdots \cup C_n)$ can be written as a sum of the type described in the statement of the lemma. Suppose that $x \in \mathrm{co}(C_1 \cup \cdots \cup C_n)$. Then there are positive numbers $s_1, \ldots, s_m$ that sum to 1 and elements $y_1, \ldots, y_m$ of $C_1 \cup \cdots \cup C_n$ such that $x = \sum_{k=1}^m s_ky_k$. For notational convenience in what is to follow, it may be assumed that the elements y_k come from more than one C_j. By reordering the indices if necessary, it may be assumed that there are integers $k_0, \ldots, k_q$ and $l_1, \ldots, l_q$ such that $0 = k_0 < \cdots < k_q = m$ and $1 \leq l_1 < \cdots < l_q \leq n$ such that $y_{k_{p-1}+1}, \ldots, y_{k_p} \in C_{l_p}$ when $1 \leq p \leq q$. For each integer p such that $1 \leq p \leq q$, let $t_p = \sum_{k=k_{p-1}+1}^{k_p} s_k$ and let $x_p = \sum_{k=k_{p-1}+1}^{k_p} t_p^{-1}s_ky_k$. Then the convexity of each C_{l_p} implies that $x_p \in C_{l_p}$ for each p. Since $\sum_{p=1}^q t_p = 1$ and $x = \sum_{p=1}^q t_px_p$, the lemma is proved. ■

2.10.14 Lemma. (N. Bourbaki, 1953 [34, p. 80]). *Suppose that $K_1, \ldots, K_n$ are compact convex subsets of a TVS. Then $\mathrm{co}(K_1 \cup \cdots \cup K_n)$ is compact.*

PROOF. Each K_j may be assumed to be nonempty. Let $\left(\sum_{j=1}^n t_j^{(\alpha)}x_j^{(\alpha)}\right)_{\alpha\in I}$ be a net in $\mathrm{co}(K_1 \cup \cdots \cup K_n)$, where each term of the net is being written as a sum of the type described in Lemma 2.10.13. Since each K_j is compact, as is the interval $[0, 1]$, there is a subnet $\left(\sum_{j=1}^n t_j^{(\beta)}x_j^{(\beta)}\right)_{\beta\in J}$ of $\left(\sum_{j=1}^n t_j^{(\alpha)}x_j^{(\alpha)}\right)_{\alpha\in I}$ such that, for each j, there are elements t_j of $[0, 1]$ and x_j of K_j such that $t_j^{(\beta)} \to t_j$ and $x_j^{(\beta)} \to x_j$. It then follows from the continuity of the vector space operations that $\sum_{j=1}^n t_j^{(\beta)}x_j^{(\beta)} \to \sum_{j=1}^n t_jx_j$. Also, the continuity of the map $(\alpha_1, \ldots, \alpha_n) \mapsto \sum_{j=1}^n \alpha_j$ from Euclidean n-space into $\mathbb{F}$ assures that $\sum_{j=1}^n t_j = 1$, so $\sum_{j=1}^n t_jx_j \in \mathrm{co}(K_1 \cup \cdots \cup K_n)$. Since every net in $\mathrm{co}(K_1 \cup \cdots \cup K_n)$ has a convergent subnet with a limit in $\mathrm{co}(K_1 \cup \cdots \cup K_n)$, the set $\mathrm{co}(K_1 \cup \cdots \cup K_n)$ is compact. ■

2.10.15 Theorem. (D. P. Milman, 1947 [170]). *Let K be a nonempty compact subset of a Hausdorff LCS such that $\overline{\mathrm{co}}(K)$ is also compact. Then every extreme point of $\overline{\mathrm{co}}(K)$ lies in K.*

PROOF. Let X be the Hausdorff LCS of which K is a subset, let x_e be an extreme point of $\overline{\mathrm{co}}(K)$, and let U be a neighborhood of 0 in X. To show that $x_e \in K$, it is enough to show that $x_e + U$ intersects K. By Theorem 2.2.9 (g) and the local convexity of X, there is a convex balanced neighborhood V of 0 such that $\overline{V} \subseteq U$. Since K is compact, there are elements $x_1, \ldots, x_n$ of K such that $K \subseteq \bigcup_{j=1}^n (x_j + V)$.

For each integer j such that $1 \le j \le n$, let $K_j = \overline{\text{co}}\Big(K \cap \big(x_j + \overline{V}\big)\Big)$. As closed subsets of the compact set $\overline{\text{co}}(K)$, the sets $K_1, \ldots, K_n$ are all compact, and so $\text{co}(K_1 \cup \cdots \cup K_n)$ is compact by Lemma 2.10.14. Since $K \subseteq K_1 \cup \cdots \cup K_n \subseteq \overline{\text{co}}(K)$, it follows that

$$\overline{\text{co}}(K) = \overline{\text{co}}(K_1 \cup \cdots \cup K_n) = \text{co}(K_1 \cup \cdots \cup K_n).$$

By Lemma 2.10.13, there are nonnegative real numbers $t_1, \ldots, t_n$ summing to 1 and elements $y_1, \ldots, y_n$ of $K_1, \ldots, K_n$ respectively such that $x_e = \sum_{j=1}^n t_j y_j$. An application of Lemma 2.10.12 yields a j_0 such that $x_e = y_{j_0}$, which implies that

$$x_e \in K_{j_0} = \overline{\text{co}}\Big(K \cap \big(x_{j_0} + \overline{V}\big)\Big) \subseteq x_{j_0} + \overline{V} = x_{j_0} - \overline{V}.$$

It follows that $x_{j_0} \in \big(x_e + \overline{V}\big) \cap K \subseteq (x_e + U) \cap K$, so $x_e + U$ intersects K. ■

2.10.16 Corollary. *Let A be a nonempty subset of a Hausdorff LCS such that $\overline{\text{co}}(A)$ is compact. Then every extreme point of $\overline{\text{co}}(A)$ lies in $\overline{A}$.*

PROOF. Since $\overline{A}$ is compact and $\overline{\text{co}}(A) = \overline{\text{co}}\big(\overline{A}\big)$, this corollary follows immediately from the theorem. ■

2.10.17 Corollary. *Let K be a nonempty weakly compact subset of a Banach space. Then every extreme point of $\overline{\text{co}}(K)$ lies in K.*

PROOF. By the Krein-Šmulian weak compactness theorem, the closed convex hull of K (which is the same as the weakly closed convex hull of K) is weakly compact, so K contains all of the extreme points of its closed convex hull. ■

2.10.18 Corollary. *Suppose that X is a normed space and that K is a nonempty subset of X^* that is bounded and weakly* closed. Then every extreme point of $\overline{\text{co}}^{w^*}(K)$ lies in K.*

PROOF. It is an easy consequence of the Banach-Alaoglu that both K and $\overline{\text{co}}^{w^*}(K)$ are weakly* compact, so the corollary follows immediately. ■

Exercises

2.107 Let x be an element of a nonempty closed convex subset C of a Hausdorff TVS. Show that x is an extreme point of C if and only if it has this property: Whenever $x_1, x_2 \in C$ and $x = \frac{1}{2}(x_1 + x_2)$, it follows that $x_1 = x_2 = x$.

2.108 Identify all of the extreme points of the closed unit ball of ℓ_1, then show that B_{ℓ_1} is the closed convex hull of its set of extreme points. Do not use any results from this section in your arguments.

2.109 Identify all of the extreme points of the closed unit ball of ℓ_∞, then show that B_{ℓ_∞} is the closed convex hull of its set of extreme points. Do not use any results from this section in your arguments.

2.110 Prove that $B_{L_1[0,1]}$ has no extreme points.

2.111 Prove that real $C[0,1]$ is not isometrically isomorphic to the dual space of any normed space.

2.112 Let C be a nonempty closed convex subset of a Hausdorff TVS X. An element x_e of C is an *exposed point* of C if there is an x^* in X^* such that $\operatorname{Re} x^*$ is bounded from above on C and attains its supremum on C at x_e and only at x_e.

(a) Show that an exposed point of C must be an extreme point of C.

(b) Find a Banach space X_0 such that some extreme point of B_{X_0} is not an exposed point of B_{X_0}. (This can be done by constructing an appropriate norm on $\mathbb{R}^2$. Exercise 1.37 might be helpful.)

Exposed points were introduced by S. Straszewicz [229] in 1935. Two good sources for more on exposed points and related objects are Klee's 1958 paper [136] and Day's book [56, pp. 105–106].

2.113 Let P be a polygon in the Euclidean plane, not assumed to be convex, whose sides intersect only at its vertices and whose vertices are each common to exactly two sides, and let K be the compact region in the plane consisting of P together with the bounded component of its complement. In the discussion of such polygonal regions that precedes Lemma 2.10.12, it is claimed that "$\overline{\text{co}}(K)$ (which is the same as $\text{co}(K)$ in this case) is a closed convex polygonal region whose vertices are a subset of the vertices of K." Prove this. Base your arguments on elementary principles, without using the Krein-Milman theorem or Theorem 2.10.15 in any way.

2.114 Let A be a nonempty subset of a Hausdorff LCS X such that $\overline{\text{co}}(A)$ is compact, let E be the set of extreme points of $\overline{\text{co}}(A)$, and let B be a subset of A. Show that the following are equivalent.

(a) $\overline{\text{co}}(B) = \overline{\text{co}}(A)$.

(b) $E \subseteq \overline{B}$.

(c) $\sup\{\operatorname{Re} x^*x : x \in B\} = \sup\{\operatorname{Re} x^*x : x \in A\}$ whenever $x^* \in X^*$.

Do not overlook the possibility that B might be empty.

*2.11 Support Points and Subreflexivity

The ideas and results of this section have close ties to James's theorem and James's weak compactness theorem from Sections 1.13 and 2.9, but do not actually depend on the material of those or any other optional sections of this book.

Recall that a bounded linear functional x^* on a normed space X is said to be *norm-attaining* if there is an x_0 in B_X such that

$$|x^*x_0| = \sup\{\, |x^*x| : x \in B_X\} = \|x^*\|.$$

As was shown in Section 1.13 and again in Section 2.9, a Banach space is reflexive if and only if each bounded linear functional on the space is norm-attaining. About the time that Robert C. James's proof of that fact for separable Banach spaces appeared in 1957 [111], Robert R. Phelps began investigating norm-attaining bounded linear functionals on nonreflexive Banach spaces, and was struck by the fact that each classical nonreflexive Banach space X has the property that the collection of norm-attaining members of X^* is dense in X^*. In light of James's work, it seemed to Phelps that a normed space X such that the norm-attaining members of X^* are dense in X^* is in a sense almost reflexive, which is the reason he gave such spaces the following name.

2.11.1 Definition. (R. R. Phelps, 1957 [184]). A normed space X is *subreflexive* if the set of norm-attaining members of X^* is dense in X^*.

Early in his study of subreflexivity, Phelps established that not every incomplete normed space is subreflexive, giving an example of a nonsubreflexive one in his 1957 paper [184] that introduced subreflexivity. The following somewhat simpler example is attributed to Yitzhak Katznelson in [28].

2.11.2 Example. Let X be the subspace of real $C[0,1]$ consisting of the polynomials on $[0,1]$. Since the Weierstrass approximation theorem assures that X is dense in $C[0,1]$, the dual space of X can be identified with $\mathrm{rca}[0,1]$ in the same way as is the dual space of $C[0,1]$. For the rest of this example, consider X^* to be so identified. The collection of all possible norm-attaining members of X^* will now be found.

Suppose that p is a constant polynomial in S_X, that is, that p is either identically 1 or identically -1 on $[0,1]$. Let μ be a member of X^* such that $|\mu(p)| = \|\mu\|$, and let μ_+ and μ_- be the positive and negative variations of μ respectively, so that $\mu = \mu_+ - \mu_-$ and $|\mu| = \mu_+ + \mu_-$. Then

$$\begin{aligned}|\mu_+([0,1]) - \mu_-([0,1])| &= |\mu([0,1])| \\ &= \left|\int_{[0,1]} p\,d\mu\right| \\ &= \|\mu\| \\ &= \mu_+([0,1]) + \mu_-([0,1]),\end{aligned}$$

from which it follows that either μ_+ or μ_- is the zero measure, that is, that μ is either a nonnegative or a nonpositive measure.

Now suppose that p is a nonconstant member of S_X and that F is the nonempty finite subset of $[0,1]$ on which $|p|$ is 1. Let μ be a member of X^* such that $|\mu(p)| = \|\mu\|$. If $|\mu|([0,1] \setminus F) > 0$, then

$$\begin{aligned}\|\mu\| &= \left|\int_{[0,1]} p\,d\mu\right| \\ &\le \int_{[0,1]\setminus F} |p|\,d|\mu| + \int_F d|\mu| \\ &< \int_{[0,1]\setminus F} d|\mu| + \int_F d|\mu| \\ &= \|\mu\|,\end{aligned}$$

a contradiction that shows that $|\mu|([0,1] \setminus F) = 0$, that is, that μ is *finitely supported.*

Thus, a member of X^* that is norm-attaining must be nonnegative, nonpositive, or finitely supported. Let λ be Lebesgue measure on $[0,1]$ and let μ_0 be the member of X^* given by the formula

$$\mu_0(A) = \lambda\big(A \cap [0,1/2]\big) - \lambda\big(A \cap [1/2,1]\big).$$

Then $\|\mu_0\| = 1$. If μ is a member of X^* that is finitely supported, then it is easy to see that $\|\mu - \mu_0\| = \|\mu\| + \|\mu_0\| \ge 1$, whereas if μ is a member of X^* that is nonnegative or nonpositive, then it is equally easy to see that $\|\mu - \mu_0\| \ge 1/2$. It follows that the open ball of radius $1/2$ centered at μ_0 contains no norm-attaining members of X^*, and therefore that the collection of norm-attaining members of X^* is not dense in X^*.

It is possible for an incomplete normed space to be subreflexive; see Exercise 2.115. In fact, as was shown by an example due to James [116], it is even possible for *all* of the bounded linear functionals on an incomplete normed space to be norm-attaining.

Examples such as the one given above leave open the possibility that every *Banach* space is subreflexive, and Phelps's early work on subreflexivity was aimed toward finding a proof of this. After Phelps discussed a new approach toward the problem with Errett Bishop, the two were able to devise a proof that appeared in 1961 [28]. The proof to be given here is based on a later reformulation of their technique in terms of *support cones* that allowed Bishop and Phelps to obtain much more general results about support points and support functionals of nonempty closed convex subsets of Banach spaces. The notation and arguments used here will closely follow those of Bishop and Phelps.

2.11.3 Definition. Let A be a subset of a topological vector space X. A nonzero x^* in X^* is a *support functional* for A if there is an x_0 in A such that $\operatorname{Re} x^*x_0 = \sup\{\operatorname{Re} x^*x : x \in A\}$, in which case x_0 is a *support point* of A and x^* *supports* A at x_0.

Support points have been encountered before in disguise, for Corollary 1.9.8 is just a roundabout way of saying that if X is any normed space, then every point of S_X is a support point of B_X.

The notion of a support functional is in a sense a generalization of the notion of a norm-attaining functional, for if x^* is a nonzero member of the dual space of a normed space X, then it follows easily from the fact that $\sup\{\operatorname{Re} x^*x : x \in B_X\} = \sup\{|x^*x| : x \in B_X\}$ that x^* is norm-attaining if and only if x^* supports B_X.

If a subset A of a TVS X is supported at some point x_0 by an x^* in X^* and U is a neighborhood of x_0, then the continuity of $\operatorname{Re} x^*$ implies that there is a y in U such that $\operatorname{Re} x^*y > \sup\{\operatorname{Re} x^*x : x \in A\}$, which implies that $y \notin A$. It follows that *every support point of a subset of a TVS is a boundary point of that set.*

2.11.4 Definition. Let X be a vector space. A subset K of X is a *wedge* if it is nonempty, convex, and has the property that $tK \subseteq K$ whenever $t \geq 0$. A wedge K in X is a *cone* if $K \cap (-K) = \{0\}$.

See Exercise 2.117 for other equivalent definitions of wedges and cones that are sometimes used. Notice that if K is a wedge, then $0K \subseteq K$, so wedges always contain 0.

2.11.5 Definition. (E. Bishop and R. R. Phelps, 1963 [29]). Suppose that X is a TVS such that $X \neq \{0\}$. Suppose further that x_0 is an element of a subset A of X and that K is a cone in X with nonempty interior such that $A \cap (x_0 + K) = \{x_0\}$. Then K is a *support cone* for A, the point x_0 is a *conical support point* of A, and $x_0 + K$ *supports* A at x_0.

The reason for defining conical support points only in nontrivial topological vector spaces is that without this restriction, the convex subset $\{0\}$ of the normed space $\{0\}$ would have 0 as a conical support point but not as a support point. This would be inconsistent with the behavior of conical support points of convex sets in larger topological vector spaces, as the following proposition shows.

2.11.6 Proposition. (E. Bishop and R. R. Phelps, 1963 [29]). *Suppose that X is a TVS such that $X \neq \{0\}$. Then every conical support point of a convex subset of X is a support point of that set.*

PROOF. Suppose that x_0 is a conical support point of a convex subset C of X. Let K be a cone in X with nonempty interior such that $x_0 + K$ supports C at x_0. If $0 \in K^\circ$, then the continuity of multiplication of vectors of X by scalars and the fact that X contains nonzero vectors together assure that there is a nonzero y_0 in X such that y_0 and $-y_0$ are both in K, which contradicts the fact that $K \cap (-K) = \{0\}$. It follows that $x_0 \notin (x_0 + K)^\circ$, and therefore that $C \cap (x_0 + K)^\circ = \varnothing$. By Eidelheit's separation theorem, there is an x^* in X^* and a real number s such that $\operatorname{Re} x^*x \le s$ for each x in C, $\operatorname{Re} x^*x \ge s$ for each x in $x_0 + K$, and $\operatorname{Re} x^*x > s$ for each x in $(x_0+K)^\circ$. It follows that $x^* \neq 0$ and that $\operatorname{Re} x^*x_0 = s = \sup\{ \operatorname{Re} x^*x : x \in C \}$, and therefore that x^* supports C at x_0. ■

Some temporary notation is now needed that will not apply outside this section. Suppose that X is a normed space, that $x^* \in S_{X^*}$, and that $t > 1$. Then

$$K(x^*, t) = \{ x : x \in X,\ \|x\| \le \operatorname{Re} x^*(tx) \}.$$

Since $\|x^*\| = 1$, there is an x_0 in X such that $\operatorname{Re} x^*(tx_0) > \|x_0\|$, which assures that $\{ x : x \in X,\ \|x\| < \operatorname{Re} x^*(tx) \}$ is a nonempty open set inside $K(x^*, t)$. It then follows easily that $K(x^*, t)$ is a closed cone with nonempty interior.

2.11.7 Lemma. (E. Bishop and R. R. Phelps, 1963 [29]). *Suppose that x is an element of a complete subset A of a normed space X, that x^* is a member of S_{X^*} such that $\operatorname{Re} x^*$ is bounded from above on A, and that $t > 1$. Then there is an x_0 in A such that $x_0 \in x + K(x^*, t)$ and $x_0 + K(x^*, t)$ supports A at x_0.*

PROOF. Let $B = A \cap \big(x + K(x^*, t)\big)$. Notice that an element z of A is in B if and only if $\|z - x\| \le \operatorname{Re} x^*\big(t(z - x)\big)$. The set B is a closed subset of the complete set A and so is itself complete. Since $K(x^*, t)$ is a cone, it follows easily that the relation on B defined by declaring that $y \preceq z$ when $z - y \in K(x^*, t)$ is a partial order for B. It is also clear that $x \in B$ and that $x \preceq z$ whenever $z \in B$.

Suppose that $\mathfrak{C}$ is a nonempty chain in B. Then $\mathfrak{C}$ can be used as the index set for a net, so the restriction of $\operatorname{Re} x^*$ to $\mathfrak{C}$ is a net in $\mathbb{R}$. Since $\operatorname{Re} x^*y \le \operatorname{Re} x^*z$ whenever $y \preceq z$, the boundedness from above of $\operatorname{Re} x^*$ on A implies that the net $\operatorname{Re} x^*|_{\mathfrak{C}}$ converges and therefore is Cauchy. Since $\|z - y\| \le \operatorname{Re} x^*\big(t(z - y)\big)$ whenever $y, z \in B$ and $y \preceq z$, the net $\mathfrak{C}$ (or, more properly, the identity function on $\mathfrak{C}$, viewed as a net) is a Cauchy net in B and so converges to some u in B. Since $K(x^*, t)$ is closed, it follows that $u - y \in K(x^*, t)$ for each y in $\mathfrak{C}$, that is, that u is an upper bound for $\mathfrak{C}$ in B.

By Zorn's lemma, the set B contains a maximal element x_0. Then $x_0 \in x + K(x^*, t)$ since $x \preceq x_0$. Suppose that $y \in A \cap \big(x_0 + K(x^*, t)\big)$. Then

$$y \in x_0 + K(x^*, t) \subseteq x + K(x^*, t) + K(x^*, t) = x + K(x^*, t),$$

so $y \in B$ and $x_0 \preceq y$. It follows from the maximality of x_0 that $y = x_0$, and therefore that $A \cap \big(x_0 + K(x^*, t)\big) = \{x_0\}$. The set $x_0 + K(x^*, t)$ therefore supports A at x_0. ■

2.11.8 Theorem. (E. Bishop and R. R. Phelps, 1963 [29]). *If C is a closed convex subset of a Banach space X such that $X \neq \{0\}$, then the conical support points of C are dense in the boundary of C.*

PROOF. Suppose that $x \in \partial C$ and that $\epsilon > 0$. Let y_0 be an element of $X \setminus C$ such that $\|y_0 - x\| < \epsilon/2$. By Theorem 2.2.28, there is an x^* in X^* such that $\sup\{\operatorname{Re} x^* y : y \in C\} < \operatorname{Re} x^* y_0$; it may be assumed that $\|x^*\| = 1$. An application of Lemma 2.11.7 produces an x_0 in C such that $x_0 \in x + K(x^*, 2)$ and $x_0 + K(x^*, 2)$ supports C at x_0. Since $x_0 - x \in K(x^*, 2)$ and $x_0 \in C$,

$$\|x_0 - x\| \leq 2\operatorname{Re} x^*(x_0 - x) < 2\operatorname{Re} x^*(y_0 - x) \leq 2\,\|y_0 - x\| < \epsilon;$$

that is, the conical support point x_0 of C, which is in ∂C by Proposition 2.11.6 and the fact that the support points of C all lie in ∂C, is less than distance ϵ from x. ■

The preceding theorem, with a little help from Proposition 2.11.6, immediately yields the following result when the Banach space is not $\{0\}$. If the space is $\{0\}$, then the result is trivially true since both subsets of the space have empty boundaries.

2.11.9 The Bishop-Phelps Support Point Theorem. (E. Bishop and R. R. Phelps, 1963 [29]). *If C is a closed convex subset of a Banach space, then the support points of C are dense in the boundary of C.*

In 1958, Victor Klee [136] asked if a nonempty bounded closed convex subset of a Banach space X such that $X \neq \{0\}$ has to have a support point. Every such set, and in fact every nonempty proper subset A of a TVS X, has a nonempty boundary; fix an x in A and a y in $X \setminus A$, let

$$s = \sup\{\, t : 0 \leq t \leq 1,\ (1-t)x + ty \in A \,\},$$

and consider $(1-s)x+sy$. The Bishop-Phelps support point theorem therefore implies that the answer to Klee's question is yes.

Let C be a closed convex subset of a Banach space X. Then in addition to the support points of C being dense in the boundary of C, it turns out that the support functionals for C must be dense in X^* provided that a few restrictions are placed on C and X. To see what those restrictions might be, first notice that C cannot be empty or all of X, for then C would have no support functionals at all. It is not enough of a restriction on C just to require that it be nonempty and not all of X; see Exercise 2.119. As will be shown, it *is* enough to require C to be nonempty and bounded, provided that $X \neq \{0\}$ so that C cannot be X. This will follow easily from a separation theorem due to Bishop and Phelps. The proof of this separation theorem is based on Lemma 2.11.11 below, which is in turn a consequence of the following result, which says, roughly speaking, that if the kernels of two norm-one bounded linear functionals on a real normed space are about the same near the origin, then each functional is about the same as the other or the negative of the other.

2.11.10 Lemma. (R. R. Phelps, 1960 [186]). *Suppose that X is a real normed space, that $\epsilon > 0$, and that x_1^* and x_2^* are members of S_{X^*} such that $x_2^*(B_X \cap \ker x_1^*) \subseteq [-\epsilon/2, \epsilon/2]$. Then either $\|x_1^* - x_2^*\| \leq \epsilon$ or $\|x_1^* + x_2^*\| \leq \epsilon$.*

PROOF. The normed space version of the Hahn-Banach extension theorem produces a y^* in X^* such that the restrictions of y^* and x_2^* to $\ker x_1^*$ are the same and $\|y^*\| \leq \epsilon/2$. Since $\ker(x_2^* - y^*) \supseteq \ker x_1^*$, it follows from Lemma 1.9.11 that there is a scalar α such that $x_2^* - y^* = \alpha x_1^*$, and

$$\big|1 - |\alpha|\big| = \big|\, \|x_2^*\| - \|\alpha x_1^*\| \,\big| \leq \|x_2^* - \alpha x_1^*\| \leq \frac{\epsilon}{2}.$$

If $\alpha \geq 0$, then

$$\|x_2^* - x_1^*\| \leq \|x_2^* - \alpha x_1^*\| + |1 - \alpha| \, \|x_1^*\| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

whereas if $\alpha < 0$, then

$$\|x_2^* + x_1^*\| \leq \|x_2^* - \alpha x_1^*\| + |1 + \alpha| \, \|x_1^*\| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

In either case, the conclusion of the lemma holds. ■

2.11.11 Lemma. (E. Bishop and R. R. Phelps, 1963 [29]). *Suppose that X is a real normed space, that $x_1^*, x_2^* \in S_{X^*}$, that $0 < \epsilon < 1$, that $t > 1 + 2/\epsilon$, and that x_2^* is nonnegative on $K(x_1^*, t)$. Then $\|x_1^* - x_2^*\| \leq \epsilon$.*

PROOF. Fix an x_0 in S_X such that $x_1^* x_0 > \max\{t^{-1}(1 + 2/\epsilon), \epsilon\}$; notice that this implies that $\|x_0\| = 1 < t x_1^* x_0$, so $x_0 \in K(x_1^*, t)$ and $x_2^* x_0 \geq 0$. If $x \in B_X \cap \ker x_1^*$, then

$$\left\| x_0 \pm \frac{2}{\epsilon} x \right\| \leq 1 + \frac{2}{\epsilon} < t x_1^* x_0 = t x_1^* \left(x_0 \pm \frac{2}{\epsilon} x \right),$$

so $x_0 \pm (2/\epsilon)x \in K(x_1^*, t)$ and therefore $x_2^*(x_0 \pm (2/\epsilon)x) \geq 0$, from which it follows that

$$|x_2^* x| = \frac{\epsilon}{2} \left| x_2^* \left(\frac{2}{\epsilon} x \right) \right| \leq \frac{\epsilon}{2} x_2^* x_0 \leq \frac{\epsilon}{2}.$$

By Lemma 2.11.10, either $\|x_1^* - x_2^*\| \leq \epsilon$ or $\|x_1^* + x_2^*\| \leq \epsilon$. However, the second inequality cannot hold, since $\|x_1^* + x_2^*\| \geq (x_1^* + x_2^*)(x_0) > \epsilon$, which completes the proof. ■

The following theorem says that, under the hypotheses on X, C, and A, any member of S_{X^*} that strictly separates C from A can be approximated by members of S_{X^*} that support C and strictly separate C from A.

2.11.12 Theorem. (E. Bishop and R. R. Phelps, 1963 [29]). *Suppose that C and A are nonempty subsets of a Banach space X such that C is closed and convex and A is bounded. Suppose further that $\epsilon > 0$, that $x_1^* \in S_{X^*}$, and that*

$$\sup\{\operatorname{Re} x_1^* x : x \in C\} < \inf\{\operatorname{Re} x_1^* x : x \in A\}.$$

Then there is an x_2^* in S_{X^*} and an x_0 in C such that

$$\operatorname{Re} x_2^* x_0 = \sup\{\operatorname{Re} x_2^* x : x \in C\} < \inf\{\operatorname{Re} x_2^* x : x \in A\}$$

and $\|x_1^* - x_2^*\| \le \epsilon$.

PROOF. It may be assumed that $0 < \epsilon < 1$ and that X is a real Banach space. Let

$$\mu = \frac{1}{2}(\inf\{x_1^* x : x \in A\} + \sup\{x_1^* x : x \in C\})$$

and

$$\delta = \frac{1}{2}(\inf\{x_1^* x : x \in A\} - \sup\{x_1^* x : x \in C\}).$$

Let $B = A + \delta B_X$, a bounded set that includes A in its interior. Then

$$\begin{aligned}\inf\{x_1^* x : x \in B\} &= \inf\{x_1^* x : x \in A\} + \delta \inf\{x_1^* x : x \in B_X\}\\ &= \inf\{x_1^* x : x \in A\} - \delta\\ &= \mu.\end{aligned}$$

Let $s = 1 + 2/\epsilon$ and select a z from C such that

$$\sup\{x_1^* x : x \in C\} - x_1^* z < \frac{\delta}{2s}.$$

Let M be any number larger than both $\delta/2$ and $\sup\{\|y - z\| : y \in B\}$ and let $t = 2sM/\delta$; then $t > s > 1$. By Lemma 2.11.7, there is an x_0 in C such that $x_0 - z \in K(x_1^*, t)$ and $x_0 + K(x_1^*, t)$ supports C at x_0.

Suppose that $y \in B$. Then

$$\begin{aligned}\|y - x_0\| &\le \|y - z\| + \|x_0 - z\|\\ &< M + x_1^*(t(x_0 - z))\\ &\le M + t(\sup\{x_1^* x : x \in C\} - x_1^* z)\\ &< M + \frac{t\delta}{2s}\\ &= \frac{t\delta}{s}\\ &< t\delta\\ &\le t x_1^*(y - x_0).\end{aligned}$$

Therefore $y - x_0 \in K(x_1^*, t)$, and so $B \subseteq x_0 + K(x_1^*, t)$.

An application of Eidelheit's separation theorem yields an x_2^* in S_{X^*} such that

$$\begin{aligned} x_2^* x_0 &= \sup\{\, x_2^* x : x \in C \,\} \\ &= \inf\{\, x_2^* x : x \in x_0 + K(x_1^*, t) \,\} \\ &\le \inf\{\, x_2^* x : x \in B \,\} \\ &= \inf\{\, x_2^* x : x \in A \,\} + \delta \inf\{\, x_2^* x : x \in B_X \} \\ &= \inf\{\, x_2^* x : x \in A \,\} - \delta \\ &< \inf\{\, x_2^* x : x \in A \,\}. \end{aligned}$$

Since $t > 1 + 2/\epsilon$ and

$$\inf\{\, x_2^* x : x \in K(x_1^*, t) \,\} = \inf\{\, x_2^* x : x \in x_0 + K(x_1^*, t) \,\} - x_2^* x_0 = 0,$$

it follows from Lemma 2.11.11 that $\|x_1^* - x_2\| \le \epsilon$. ■

Two of the main results of this section are easy consequences of the theorem just proved.

2.11.13 The Bishop-Phelps Support Functional Theorem. (E. Bishop and R. R. Phelps, 1963 [29]). *If C is a nonempty bounded closed convex subset of a Banach space X such that $X \neq \{0\}$, then the support functionals for C are dense in X^*.*

PROOF. It suffices to prove that the norm-one support functionals for C are dense in S_{X^*}. Let x_1^* be an element of S_{X^*}, let y be an element of X such that $\sup\{\, \mathrm{Re}\, x_1^* x : x \in C \,\} < \mathrm{Re}\, x_1^* y$, and let $\epsilon > 0$. Letting $A = \{y\}$ in Theorem 2.11.12 yields a norm-one support functional x_2^* for C such that $\|x_1^* - x_2^*\| \le \epsilon$. ■

The completeness hypothesis in the preceding result is necessary, for Bishop and Phelps proved in [29] that every incomplete normed space X has a bounded closed convex subset C with nonempty interior such that the support functionals for C are not dense in X^*.

Applying the preceding result to the closed unit ball of a Banach space immediately yields the subreflexivity result that motivated this entire section. (The special case in which the Banach space is $\{0\}$ is trivial.)

2.11.14 The Bishop-Phelps Subreflexivity Theorem. (E. Bishop and R. R. Phelps, 1961 [28]). *Every Banach space is subreflexive.*

Suppose that X and Y are normed spaces. As one would expect, a member T of $B(X, Y)$ is called *norm-attaining* if there is an x in B_X such that $\|Tx\| = \|T\|$. In light of the Bishop-Phelps subreflexivity theorem, it is natural to ask the following question that Bishop and Phelps posed in [28]: *What conditions on two Banach spaces X and Y assure that the collection of norm-attaining members of $B(X,Y)$ is dense in $B(X,Y)$?* There must be some additional conditions imposed on at least one of the spaces; in particular, in a 1963 paper Joram Lindenstrauss [153] showed that there is a Banach space X such that the set of norm-attainers in $B(X)$ is not dense in $B(X)$. Partial answers to the question have emerged from research done on a property for Banach spaces called the *Radon-Nikodým property,* in particular from work by Jean Bourgain [35] published in 1977. See the books by Richard Bourgin [37] and by Joseph Diestel and J. Jerry Uhl [59] for discussions of the Radon-Nikodým property and its relationship to the question raised by Bishop and Phelps.

Since Bishop and Phelps showed that the norm-attainers are dense in $B(X,Y)$ for every Banach space X when $Y = \mathbb{F}$, it is particularly natural to ask what conditions on a Banach space Y would assure that the norm-attainers are dense in $B(X,Y)$ for every Banach space X. A good starting point for anyone interested in that problem is Timothy Gowers's 1990 paper [90], in which he shows that each space ℓ_p such that $1 < p < \infty$ lacks this property.

Exercises

2.115 Let X be the vector space of finitely nonzero sequences, equipped with the ℓ_2 norm. Show that X is an incomplete subreflexive normed space.

2.116 Suppose that C is a closed convex subset of a TVS such that $C^\circ \neq \varnothing$. Prove that every point of the boundary of C is a support point of C.

2.117 Let K be a nonempty subset of a vector space X.

(a) Prove that K is a wedge if and only if it has this property: If $x, y \in K$ and $t \geq 0$, then $x + y, tx \in K$. (Notice the natural way in which wedges are generalizations of subspaces.)

(b) Prove that K is a wedge if and only if it has this property: If $x, y \in K$ and $s, t \geq 0$, then $sx + ty \in K$. (Notice the natural way in which convex sets are generalizations of wedges.)

(c) Prove that K is a cone if and only if it is a wedge with the property that $x = 0$ whenever $x, -x \in K$.

2.118 The preceding exercise is useful, though not at all essential, for this one. Suppose that P is a wedge in a real vector space X. Define a relation $\preceq$ on X by declaring that $x \preceq y$ when $y - x \in P$.

(a) Prove that $\preceq$ is a preorder on X.

(b) Prove that if $x, y, z \in X$, $t \geq 0$, and $x \preceq y$, then $x + z \preceq y + z$ and $tx \preceq ty$.

(c) Prove that $\preceq$ is a partial order if and only if P is a cone.

The pair $(X, \preceq)$ is called an *ordered vector space,* the members of P are called the *positive elements* of X, and P itself is called the *positive wedge* (or *positive cone,* if appropriate) of X.

2.119 Find a closed convex subset C of a Banach space X such that the collection of support functionals for C is nonempty but not dense in X^*. (It can be done in $\mathbb{R}^2$.)

2.120 Call a subset A of a normed space X *orthodox* if for each x in $X \setminus A$ there is a norm-attaining x^* in X^* such that $\operatorname{Re} x^*x > \sup\{\operatorname{Re} x^*y : y \in A\}$. Prove that a normed space is subreflexive if and only if each of its bounded closed convex subsets is orthodox.

2.121 Prove the following improvement of Theorem 2.2.28 for Banach spaces: *Let K and C be disjoint nonempty convex subsets of a Banach space X such that K is compact and C is closed. Then there is a member x^* of X^* such that* $\inf\{\operatorname{Re} x^*x : x \in C\}$ *is attained by* $\operatorname{Re} x^*$ *on C and* $\max\{\operatorname{Re} x^*x : x \in K\} < \min\{\operatorname{Re} x^*x : x \in C\}$.

2.122 Let X be a normed space such that $X \neq \{0\}$. If $t \in \mathbb{R}$ and x^* is a nonzero member of X^*, then the subset of X defined by the formula

$$H(x^*, t) = \{x : x \in X,\ \operatorname{Re} x^*x \leq t\}$$

is called a *closed halfspace.* Let C be a nonempty closed convex subset of X. Then a closed halfspace $H(x^*, t)$ is said to *support* C if $t = \sup\{\operatorname{Re} x^*x : x \in C\}$ and x^* is a supporting functional for C. It was shown in Exercise 1.109 (which is not needed for this exercise) that C is the intersection of a collection of closed halfspaces. Prove that if X is a Banach space, then C is the intersection of the collection of its supporting closed halfspaces.

2.123 Suppose that A is a subset of a topological vector space X. A nonzero x^* in X^* is a *modulus support functional* for A if there is an x_0 in A such that $|x^*x_0| = \sup\{|x^*x| : x \in A\}$, in which case x_0 is said to be a *modulus support point* of A.

(a) Show that if A is balanced, then the modulus support points of A are the same as the support points of A, and the modulus support functionals for A are the same as the support functionals for A.

(b) Give an example of a subset C of a Banach space such that the modulus support points of C are not dense in the boundary of C, even though C is bounded, closed, convex, and has nonempty interior. (This can be done in $\mathbb{C}$.)

It turns out not to be too hard to show that every nonempty bounded closed convex subset of a real Banach space X has the property that its modulus support functionals are dense in X^*. At the time of this writing, it is not known if this result extends to complex Banach spaces. For more information, see Phelps's 1992 survey article [187] on this problem.

2.124 Let X be a normed space. By Corollary 1.9.8, every member of S_X is a support point of B_X. The space X is said to be *smooth* if for each x

in S_X there is a unique x_x^* in S_{X^*} that supports B_X at x. The space X is then said to be *very smooth* if it is smooth and the map $x \mapsto x_x^*$ from S_X into S_{X^*} is norm-to-weak continuous. These definitions will be encountered again in Chapter 5.

(a) Prove that if X is a Banach space and X^* is very smooth, then every member of X^* is norm-attaining.

(b) (This part depends on James's theorem from Section 1.13 or 2.9.) Conclude from (a) that every Banach space with a very smooth dual is reflexive.

3

Linear Operators

Linear operators have already received quite a bit of attention in this book, primarily as tools for probing the structure of normed spaces. The purpose of this chapter is to reverse that emphasis temporarily by studying linear operators between normed spaces as interesting objects in their own right, with the properties of normed spaces obtained in the first two chapters used as the tools for the study. Almost all of the attention will be on bounded linear operators, though there is also an interesting theory of unbounded ones; see, for example, Chapter 13 of [200] or Chapter VII, Section 9 of [67].

The adjoint of an isomorphism from one normed space onto another was introduced in Section 1.10, primarily to be able to prove that normed spaces that are isomorphic or isometrically isomorphic have dual spaces bearing that same relationship to each other. The first order of business is to extend the notion of adjoint operators and learn more about them.

3.1 Adjoint Operators

Recall the following notation introduced in Section 1.10: When x is an element of a vector space X and $x^\# \in X^\#$, then $x^\# x$ is often denoted by $\langle x, x^\# \rangle$, particularly in situations in which the usual notation might be visually confusing.

3.1.1 Definition. Suppose that X and Y are vector spaces and that $T \in L(X,Y)$. Then the *algebraic adjoint* of T is the linear operator $T^\#$

from $Y^\#$ into $X^\#$ given by the formula $T^\#(y^\#) = y^\# T$. That is, the algebraic adjoint $T^\#$ is defined by letting $\langle x, T^\# y^\# \rangle = \langle Tx, y^\# \rangle$ whenever $x \in X$ and $y^\# \in Y^\#$.

With X, Y, and T as in the preceding definition, it is immediate that $T^\#(y^\#)$ really is a linear functional on X whenever $y^\#$ in $Y^\#$, and it is easy to check that $T^\#$ is itself linear.

3.1.2 Proposition. *Suppose that X and Y are normed spaces and that $T \in L(X,Y)$. Then T is bounded if and only if $T^\#(Y^*) \subseteq X^*$. If T is bounded, then the restriction T^* of $T^\#$ to Y^*, viewed as a member of $L(Y^*, X^*)$, is itself bounded, and $\|T^*\| = \|T\|$.*

PROOF. The fact that T is bounded if and only if $T^\#(Y^*) \subseteq X^*$ is just a restatement of Proposition 2.5.10. Now suppose that T is bounded, so that $T^\#(Y^*) \subseteq X^*$, and let the restriction T^* of $T^\#$ to Y^* be viewed as a member of $L(Y^*, X^*)$. An application of Proposition 1.10.11 (a) shows that

$$\begin{aligned} \|T\| &= \sup\{\, |\langle Tx, y^* \rangle| : x \in B_X,\ y^* \in B_{Y^*} \} \\ &= \sup\{\, |\langle x, T^* y^* \rangle| : x \in B_X,\ y^* \in B_{Y^*} \}. \end{aligned}$$

By Proposition 1.10.11 (b), the fact that the second supremum in the above equality is finite assures that T^* is bounded and has its norm equal to that supremum, and therefore to $\|T\|$. ■

3.1.3 Definition. Suppose that X and Y are normed spaces and that $T \in B(X,Y)$. Then the *(normed-space) adjoint* of T is the restriction of $T^\#$ to Y^*, that is, the linear operator T^* from Y^* to X^* given by the formula $T^*(y^*) = y^* T$.

Thus, with all notation as in the preceding definition, the normed-space adjoint T^* of T is defined by letting $\langle x, T^* y^* \rangle = \langle Tx, y^* \rangle$ whenever $x \in X$ and $y^* \in Y^*$. For the rest of this book, the only real interest in adjoints will be in the normed-space variety, so the unqualified term *adjoint* will always refer to them and not to algebraic adjoints.

Definition 3.1.3 is due to Banach [11] for the case in which X and Y are arbitrary Banach spaces, but adjoints of operators on special spaces were used long before the appearance of Banach's paper in 1929. In particular, Frédéric Riesz made use of adjoints in published works appearing in 1910 [193] and 1913 [194] concerning linear operators on ℓ_2 and certain other L_p spaces. The idea of an adjoint actually has its roots in matrix theory, as will now be shown.

For the purposes of the next paragraph, consider Euclidean p-space ($p > 0$) to be identified with the collection of all $p \times 1$ matrices of scalars in the obvious way, and $B(\mathbb{F}^p, \mathbb{F}^q)$ $(p, q > 0)$ to be identified with the collection of all $p \times q$ matrices of scalars, with the action of a member T of $B(\mathbb{F}^p, \mathbb{F}^q)$ on an element x of $\mathbb{F}^p$ given by the formula $Tx = T^t \cdot x$, where T^t is the transpose of the matrix T and $\cdot$ represents matrix multiplication. In particular, when $\mathbb{F}$ is identified with $\mathbb{F}^1$ in the obvious way, a member x^* of $(\mathbb{F}^p)^*$ acts on a member x of $\mathbb{F}^p$ through the formula $x^*x = (x^*)^t \cdot x$.

Suppose that m and n are positive integers and that $T \in B(\mathbb{F}^m, \mathbb{F}^n)$. If $x \in \mathbb{F}^m$ and $y^* \in (\mathbb{F}^n)^*$, then

$$\langle x, T^*y^* \rangle = \langle Tx, y^* \rangle = (y^*)^t \cdot (T^t \cdot x) = \left((y^*)^t \cdot T^t\right) \cdot x = \left((T^t)^t \cdot y^*\right)^t \cdot x,$$

so $T^*y^* = (T^t)^t \cdot y^*$. It follows that T^* is, as a matrix, equal to T^t. The notion of an adjoint operator is therefore, in a sense, a generalization of the notion of the transpose of a matrix.

The verification of the following proposition is easy. The corollary then follows after a glance at Proposition 3.1.2.

3.1.4 Proposition. *If S and T are bounded linear operators from a normed space X into a normed space Y and $\alpha \in \mathbb{F}$, then $(S+T)^* = S^* + T^*$ and $(\alpha S)^* = \alpha(S^*)$.*

3.1.5 Corollary. *If X and Y are normed spaces, then the map $T \mapsto T^*$ is an isometric isomorphism from $B(X,Y)$ into $B(Y^*, X^*)$.*

The isometric isomorphism of the preceding corollary does not have to map $B(X,Y)$ onto $B(Y^*, X^*)$, even when both X and Y are Banach spaces. See Exercises 3.3 and 3.4.

3.1.6 Example. Let I be the identity operator on a normed space X. For each x in X and each x^* in X^*,

$$\langle x, I^*x^* \rangle = \langle Ix, x^* \rangle = \langle x, x^* \rangle.$$

It follows that $I^*x^* = x^*$ for each x^* in X^*, that is, that I^* is the identity operator on X^*.

3.1.7 Example. In this example, members of ℓ_1 will also be treated as members of c_0 and ℓ_∞, so to avoid confusion a subscript of 0, 1, or ∞ will indicate whether a particular sequence of scalars is being treated as a member of c_0, ℓ_1, or ℓ_∞ respectively. Let T be the map from ℓ_1 into c_0 given by the formula $T\big((\alpha_n)_1\big) = (\alpha_n)_0$. Then T is clearly linear and bounded and is easily seen to have norm 1. Consider ℓ_1^* and c_0^* to be identified with ℓ_∞ and ℓ_1 respectively in the usual way. Then for each pair of elements $(\alpha_n)_1$ and $(\beta_n)_1$ of ℓ_1,

$$\langle (\alpha_n)_1, T^*(\beta_n)_1 \rangle = \langle T(\alpha_n)_1, (\beta_n)_1 \rangle = \langle (\alpha_n)_0, (\beta_n)_1 \rangle = \sum_n \beta_n \alpha_n,$$

so the element $T^*(\beta_n)_1$ of ℓ_1^* can be identified with the element $(\beta_n)_\infty$ of ℓ_∞. In short, the adjoint of the "identity" map from ℓ_1 into c_0 is the "identity" map from ℓ_1 into ℓ_∞. It is clear that the norm of T^* is 1, as predicted by Proposition 3.1.2.

The preceding two examples might give one the idea that the adjoint of a one-to-one bounded linear operator between normed spaces must itself be one-to-one. As the next example shows, this is not the case. The actual relationships between such properties as being one-to-one and being onto for bounded linear operators and their adjoints will be settled by Theorems 3.1.17, 3.1.18, and 3.1.22.

3.1.8 Example. Let X be any nonreflexive Banach space and let Q_X be the natural map from X into X^{**}, an isometric isomorphism from X onto a closed subspace of X^{**}. Then Q_X^* maps X^{***} into X^*. Let Q_{X^*} be the natural map from X^* into X^{***}. Then for each x in X and each x^* in X^*,

$$\langle x, Q_X^* Q_{X^*} x^* \rangle = \langle Q_X x, Q_{X^*} x^* \rangle = \langle x^*, Q_X x \rangle = \langle x, x^* \rangle,$$

which implies that $Q_X^* Q_{X^*}$ is the identity map on X^* and therefore that Q_X^* maps X^{***} onto X^*. If Q_X^* were also one-to-one, then Q_{X^*} would have to map X^* onto X^{***}, contradicting the fact that X^* is not reflexive. The isometric isomorphism Q_X therefore does not have a one-to-one adjoint.

As has already been mentioned, adjoints were used in Section 1.10 to prove that normed spaces that are isomorphic or isometrically isomorphic have dual spaces that are isomorphic or isometrically isomorphic, respectively. Adjoints had another use in that section, as the next example shows.

3.1.9 Example. Suppose that X is a normed space and that M is a closed subspace of X. By Theorem 1.10.17, there is an isometric isomorphism that identifies $(X/M)^*$ with $M^\perp$ such that if an element of $(X/M)^*$ is identified with the element x^* of $M^\perp$, then $x^*(x+M) = x^*x$ for each member $x+M$ of X/M. A glance at the proof of Theorem 1.10.17 shows that this isometric isomorphism is the adjoint of the quotient map from X onto X/M.

Since the concept of the adjoint of a linear operator is in a sense a generalization of the notion of the transpose of a matrix of scalars, it is not surprising that some of the properties of transposes of matrices generalize to adjoints of operators. In particular, the following result is a generalization of the fact that if A and B are matrices whose dimensions are such that the matrix product $A \cdot B$ is defined, then $(A \cdot B)^t = B^t \cdot A^t$.

3.1.10 Proposition. *Suppose that X, Y, and Z are normed spaces, that $S \in B(X,Y)$, and that $T \in B(Y,Z)$. Then $(TS)^* = S^*T^*$.*

PROOF. If $x \in X$ and $z^* \in Z^*$, then

$$\langle x, (TS)^*(z^*)\rangle = \langle TSx, z^*\rangle = \langle Sx, T^*z^*\rangle = \langle x, S^*T^*z^*\rangle,$$

from which it follows that $(TS)^* = S^*T^*$. ∎

It is certainly true that every matrix A is the transpose of a matrix, since $A = (A^t)^t$. There is a sense in which this fact generalizes to adjoints, as Proposition 3.1.13 and the comments immediately preceding it will show. In another sense it does not, since there exist Banach spaces X and Y such that some members of $B(Y^*, X^*)$ are not adjoints of members of $B(X,Y)$; see the comment after Corollary 3.1.5. The following result characterizes the linear operators between dual spaces of normed spaces that actually are adjoints.

3.1.11 Theorem. *Suppose that X and Y are normed spaces. If $T \in B(X,Y)$, then T^* is weak*-to-weak* continuous. Conversely, if S is a weak*-to-weak* continuous linear operator from Y^* into X^*, then there is a T in $B(X,Y)$ such that $T^* = S$.*

PROOF. Suppose first that $T \in B(X,Y)$. Let (y^*_α) be a net in Y^* that is weakly* convergent to some y^*. For each x in X,

$$\langle x, T^*y^*_\alpha\rangle = \langle Tx, y^*_\alpha\rangle \to \langle Tx, y^*\rangle = \langle x, T^*y^*\rangle,$$

so $T^*y^*_\alpha \xrightarrow{w^*} T^*y^*$. This establishes the weak*-to-weak* continuity of T^*.

Now suppose instead that S is a weak*-to-weak* continuous linear operator from Y^* into X^*. Let Q_X and Q_Y be the natural maps from X and Y respectively into their second duals. For each x in X, the weak* continuity of $Q_X(x)$ on X^* assures that the operator product $Q_X(x)S$ is a weakly* continuous linear functional on Y^* and so is a member of $Q_Y(Y)$, which in turn implies that $Q_Y^{-1}\big(Q_X(x)S\big) \in Y$.

Define $T\colon X \to Y$ by the formula $Tx = Q_Y^{-1}\big(Q_X(x)S\big)$. It is easy to check that T is linear. To see that T is bounded, suppose that (x_α) is a net in X converging weakly to some x_0. Then $Q_X(x_\alpha) \xrightarrow{w^*} Q_X(x_0)$, so $\big(Q_X(x_\alpha)S\big)(y^*) \to \big(Q_X(x_0)S\big)(y^*)$ whenever $y^* \in Y^*$, which in turn implies that $Q_X(x_\alpha)S \xrightarrow{w^*} Q_X(x_0)S$ and therefore that

$$Tx_\alpha = Q_Y^{-1}\big(Q_X(x_\alpha)S\big) \xrightarrow{w} Q_Y^{-1}\big(Q_X(x_0)S\big) = Tx_0.$$

The operator T is therefore weak-to-weak continuous and so is norm-to-norm continuous by Theorem 2.5.11.

Finally, observe that for each x in X and each y^* in Y^*,

$$\begin{aligned}\langle x, T^*y^*\rangle &= \langle Tx, y^*\rangle \\ &= \langle Q_Y^{-1}(Q_X(x)S), y^*\rangle \\ &= \langle y^*, Q_X(x)S\rangle \\ &= \langle Sy^*, Q_X(x)\rangle \\ &= \langle x, Sy^*\rangle,\end{aligned}$$

which shows that $T^* = S$. ■

3.1.12 Corollary. *Suppose that X and Y are normed spaces. Then every weak*-to-weak* continuous linear operator from X^* into Y^* is norm-to-norm continuous.*

In Example 2.6.17, an isomorphism T from c_0^* onto itself was constructed that is not weak*-to-weak* continuous, so the converse of the preceding corollary is not in general true. Notice also that, by Theorem 3.1.11, the isomorphism T is an explicit example of a member of $B(c_0^*)$ that is not the adjoint of any member of $B(c_0)$.

As was mentioned in the comments preceding Theorem 3.1.11, there is a sense in which the fact that $(A^t)^t = A$ for every matrix A has an extension to adjoints. The generalization is given in the following proposition, which says that if T is a bounded linear operator from a normed space X into a normed space Y and each space is identified with its image in its second dual under the natural map, then $T^{**} = T$ on X.

3.1.13 Proposition. *Suppose that X and Y are normed spaces and that $T \in B(X,Y)$. Let Q_X and Q_Y be the natural maps from X and Y into their respective second dual spaces. Then $T^{**}Q_X(X) \subseteq Q_Y(Y)$ and $Q_Y^{-1}T^{**}Q_X = T$.*

PROOF. Fix an x in X. First notice that whenever $y^* \in Y^*$,

$$\langle y^*, T^{**}Q_Xx\rangle = \langle T^*y^*, Q_Xx\rangle = \langle x, T^*y^*\rangle = \langle Tx, y^*\rangle.$$

If a net (y_α^*) in Y^* is weakly* convergent to some y_0^*, then

$$\langle y_\alpha^*, T^{**}Q_Xx\rangle = \langle Tx, y_\alpha^*\rangle \to \langle Tx, y_0^*\rangle = \langle y_0^*, T^{**}Q_Xx\rangle,$$

so $T^{**}Q_Xx$ is a weakly* continuous linear functional on Y^* and therefore lies in $Q_Y(Y)$. Whenever $y^* \in Y^*$,

$$\langle Tx, y^*\rangle = \langle y^*, T^{**}Q_Xx\rangle = \langle Q_Y^{-1}T^{**}Q_Xx, y^*\rangle,$$

from which it follows that $Q_Y^{-1}T^{**}Q_X = T$. ■

3.1.14 Corollary. *Suppose that X and Y are normed spaces and that $T \in B(X,Y)$. Let Q_X and Q_Y be the natural maps from X and Y into their respective second dual spaces.*

(a) *The operator T is one-to-one if and only if the restriction of T^{**} to $Q_X(X)$ is one-to-one.*

(b) *The operator T maps X onto Y if and only if T^{**} maps $Q_X(X)$ onto $Q_Y(Y)$.*

PROOF. The preceding proposition implies that $T^{**}Q_X = Q_YT$, from which the corollary follows easily. ■

Yet another property of transposes of matrices that generalizes to adjoints is the fact that a square matrix A is invertible if and only if its transpose is invertible, in which case $(A^{-1})^t = (A^t)^{-1}$. The generalization of the formula for $(A^{-1})^t$ is easy.

3.1.15 Proposition. *Suppose that T is an isomorphism from a normed space X onto a normed space Y. Then $(T^{-1})^* = (T^*)^{-1}$.*

PROOF. By Theorem 1.10.12, the operator T^* is an isomorphism from Y^* onto X^*, so $(T^*)^{-1}$ does exist. For each y in Y and each x^* in X^*,

$$\begin{aligned}\langle y, (T^{-1})^*(x^*)\rangle &= \langle T^{-1}y, x^*\rangle \\ &= \langle T^{-1}y, T^*(T^*)^{-1}(x^*)\rangle \\ &= \langle TT^{-1}y, (T^*)^{-1}(x^*)\rangle \\ &= \langle y, (T^*)^{-1}(x^*)\rangle,\end{aligned}$$

from which it follows that $(T^{-1})^* = (T^*)^{-1}$ ■

The generalization to adjoints of the fact that a square matrix is invertible if and only if its transpose is invertible requires a bit more work. One key element for obtaining the generalization is part (b) of the next theorem, The lemma used to prove the theorem is itself of some interest.

3.1.16 Lemma. *Suppose that X and Y are normed spaces and that $T \in B(X,Y)$. Then $\ker(T) = {}^{\perp}(T^*(Y^*))$ and $\ker(T^*) = \big(T(X)\big)^{\perp}$.*

PROOF. Since Y^* is a separating family of functions for Y, an element x of X is in $\ker(T)$ if and only if

$$\langle x, T^*y^*\rangle = \langle Tx, y^*\rangle = 0 \tag{3.1}$$

for each y^* in Y^*, which is the same as saying that $x \in {}^{\perp}\big(T^*(Y^*)\big)$. Similarly, an element y^* of Y^* is in $\ker(T^*)$ if and only if (3.1) holds for each x in X, which is equivalent to requiring that y^* be in $\big(T(X)\big)^{\perp}$. ■

3.1.17 Theorem. *Suppose that X and Y are normed spaces and that $T \in B(X,Y)$.*

(a) *The operator T is one-to-one if and only if $T^*(Y^*)$ is weakly* dense in X^*.*

(b) *The operator T^* is one-to-one if and only if $T(X)$ is dense in Y.*

PROOF. By Lemma 3.1.16 and Propositions 2.6.6 (c) and 1.10.15 (c),

$$\left(\ker(T)\right)^{\perp} = \left({}^{\perp}\left(T^*(Y^*)\right)\right)^{\perp} = \overline{T^*(Y^*)}^{w^*} \tag{3.2}$$

and

$${}^{\perp}\left(\ker(T^*)\right) = {}^{\perp}\left(\left(T(X)\right)^{\perp}\right) = \overline{T(X)}. \tag{3.3}$$

Since T is one-to-one if and only if $\ker(T) = \{0\}$, which happens if and only if $\left(\ker(T)\right)^{\perp} = X^*$, it follows from (3.2) that T is one-to-one if and only if $X^* = \overline{T^*(Y^*)}^{w^*}$, proving (a). Part (b) follows from a similar argument based on (3.3). ■

Here is the promised extension of the fact that square matrices are invertible exactly when their transposes are. Part of this result was previously obtained as Theorem 1.10.12. Unlike Theorem 1.10.12 and Proposition 3.1.15, this theorem has a completeness hypothesis that cannot in general be removed. See Exercise 3.8.

3.1.18 Theorem. *Suppose that X is a Banach space, that Y is a normed space, and that $T \in B(X,Y)$. Then T is an isomorphism from X onto Y if and only if T^* is an isomorphism from Y^* onto X^*. The same is true if "isomorphism" is replaced by "isometric isomorphism."*

PROOF. It has already been shown in Theorem 1.10.12 that if T is an isomorphism or isometric isomorphism from X onto Y, then T^* is an isomorphism or isometric isomorphism, respectively, from Y^* onto X^*.

Suppose that T^* is an isomorphism from Y^* onto X^*. Then there is a positive constant c such that $(T^*)^{-1}(B_{X^*}) \subseteq cB_{Y^*}$. If $x \in X$, then

$$\begin{aligned}
\|Tx\| &= \sup\{\, |\langle Tx, y^*\rangle| : y^* \in B_{Y^*}\,\} \\
&= \sup\{\, |\langle x, T^*y^*\rangle| : y^* \in B_{Y^*}\,\} \\
&\geq \sup\{\, |\langle x, x^*\rangle| : x^* \in c^{-1}B_{X^*}\,\} \\
&= c^{-1}\|x\|,
\end{aligned}$$

which is enough to assure that T is an isomorphism. Since $T(X)$ is dense in Y and complete, hence closed, it follows that $T(X) = Y$.

Suppose now that T^* is an isometric isomorphism from Y^* onto X^*. Then the constant c in the argument just given can be selected to be 1. This and the fact that $\|T\| = \|T^*\|$ assures that $\|x\| \le \|Tx\| \le \|x\|$ for each x in X, which implies that T is an isometric isomorphism. ∎

One immediate consequence of Theorem 3.1.17 is that, with all notation as in that theorem, if T maps X onto Y (respectively, T^* maps Y^* onto X^*), then T^* (respectively, T) is one-to-one. These statements can be greatly strengthened when X and Y are Banach spaces, as will be shown in Theorem 3.1.22. A major ingredient in the proof of that result is Theorem 3.1.21, which establishes important relationships between closure properties of the range of T and those of the range of T^* when X and Y are complete. The proof of Theorem 3.1.21 given here, which is essentially the one that appears in [43], uses the Krein-Šmulian theorem on weakly* closed convex sets. See [200] for a different proof that does not.

3.1.19 Lemma. *Suppose that M is a closed subspace of a normed space X and that π is the quotient map from X onto X/M. Then π^* is a homeomorphism from $(X/M)^*$ with its weak* topology onto the weakly* closed subspace $M^\perp$ of X^* with the relative weak* topology that $M^\perp$ inherits from X^*. That is, the usual identification map from $(X/M)^*$ onto $M^\perp$ also identifies the weak* topologies of the two spaces.*

PROOF. As has been noted in Example 3.1.9, the map π^* is the usual identification map from $(X/M)^*$ onto $M^\perp$ and so is one-to-one. The set $M^\perp$ is weakly* closed in X^* by Proposition 2.6.6 (a), and the weak*-to-relative-weak* continuity of π^* follows from Theorem 3.1.11, so all that needs to be shown is the relative-weak*-to-weak* continuity of $(\pi^*)^{-1}$ on $M^\perp$.

Suppose that a net (m_α^*) in $M^\perp$ is weakly* convergent to some member m^* of $M^\perp$. Then for each x in X,

$$\langle x+M, (\pi^*)^{-1}(m_\alpha^*)\rangle = m_\alpha^* x \to m^* x = \langle x+M, (\pi^*)^{-1}(m^*)\rangle,$$

so $(\pi^*)^{-1}(m_\alpha^*) \xrightarrow{w^*} (\pi^*)^{-1}(m^*)$. ∎

See Theorem 1.7.13 for the basic properties of the map S in the following lemma.

3.1.20 Lemma. *Suppose that X and Y are normed spaces and that $T \in B(X,Y)$. Let π be the quotient map from X onto $X/\ker(T)$, let S be the bounded linear operator from $X/\ker(T)$ into Y such that $T = S\pi$, let $Z = \overline{T(X)}$, and let R be T viewed as a member of $B(X,Z)$.*

(a) *The set $S(X/\ker(T))$ is closed if and only if $T(X)$ is closed.*

(b) *The set $S^*(Y^*)$ is closed if and only if $T^*(Y^*)$ is closed.*

(c) *The set* $S^*(Y^*)$ *is weakly* closed if and only if* $T^*(Y^*)$ *is weakly* closed.*

(d) *The set* $R(X)$ *is closed if and only if* $T(X)$ *is closed.*

(e) *The set* $R^*(Z^*)$ *is closed if and only if* $T^*(Y^*)$ *is closed.*

(f) *The set* $R^*(Z^*)$ *is weakly* closed if and only if* $T^*(Y^*)$ *is weakly* closed.*

PROOF. Part (a) follows immediately from the fact that the range of S is the same as that of T, while (b) and (c) hold because $T^* = \pi^* S^*$ and π^* is both an isometric isomorphism and a weak*-to-relative-weak* homeomorphism from $(X/M)^*$ onto the weakly* closed subspace $M^\perp$ of X^*. Part (d) is obvious. For (e) and (f), notice that if $y^* \in Y^*$ and z^* is the member of Z^* formed by restricting y^* to Z, then

$$\langle x, T^*y^* \rangle = \langle Tx, y^* \rangle = \langle Rx, z^* \rangle = \langle x, R^*z^* \rangle$$

whenever $x \in X$, and so $T^*y^* = R^*z^*$. Since every member of Z^* has a Hahn-Banach extension to Y and therefore is the restriction of some member of Y^* to Z, the ranges of R^* and T^* are the same, from which (e) and (f) follow immediately. ■

3.1.21 Theorem. *Suppose that* X *and* Y *are Banach spaces and that* $T \in B(X, Y)$. *Then the following are equivalent.*

(a) *The set* $T(X)$ *is closed.*

(b) *The set* $T^*(Y^*)$ *is closed.*

(c) *The set* $T^*(Y^*)$ *is weakly* closed.*

PROOF. Let π be the quotient map from X onto $X/\ker(T)$ and let S be the bounded linear operator from $X/\ker(T)$ into Y such that $T = S\pi$. Then S is one-to-one since its kernel contains only the zero element of $X/\ker(T)$. It is a consequence of this and parts (a), (b), and (c) of Lemma 3.1.20 that the theorem is true if it holds under the additional hypothesis that T is one-to-one. It may therefore be assumed that T is one-to-one. Now suppose that the theorem holds under the additional hypothesis that $\overline{T(X)} = Y$. The general case then follows from parts (d), (e), and (f) of Lemma 3.1.20, so it may also be assumed that $T(X)$ is dense in Y. By Theorem 3.1.17, the operator T^* is one-to-one and $T^*(Y^*)$ is weakly* dense in X^*.

Suppose that $T(X)$ is closed and therefore that $T(X) = Y$. Then Corollary 1.6.6 implies that T is an isomorphism from X onto Y, so it follows from Theorem 1.10.12 that T^* is an isomorphism from Y^* onto X^* and therefore that $T^*(Y^*)$ is closed. This shows that (a) $\Rightarrow$ (b).

Suppose next that $T^*(Y^*)$ is weakly* closed. Then $T^*(Y^*) = X^*$, so T^* is an isomorphism from Y^* onto X^*. It follows from Theorem 3.1.18 that T is an isomorphism from X onto Y and therefore that $T(X)$ is closed, which shows that (c) $\Rightarrow$ (a).

Finally, suppose that $T^*(Y^*)$ is closed. Since Y^* and $T^*(Y^*)$ are both Banach spaces, the operator T^* is an isomorphism from Y^* onto $T^*(Y^*)$. Suppose that (x_α^*) is a net in $T^*(Y^*) \cap B_{X^*}$ that is weakly* convergent to some x^* in X^*. Then $x^* \in B_{X^*}$ since B_{X^*} is weakly* closed. Let $y_\alpha^* = (T^*)^{-1}(x_\alpha^*)$ for each α. Then (y_α^*) is a bounded net and so, by the Banach-Alaoglu theorem, has a subnet (y_β^*) that is weakly* convergent to some y^* in Y^*. It follows from the weak*-to-weak* continuity of T^* that $x_\beta^* = T^*y_\beta^* \xrightarrow{w^*} T^*y^*$, so $T^*y^* = x^*$. Since $x^* \in T^*(Y^*) \cap B_{X^*}$, the set $T^*(Y^*) \cap B_{X^*}$ is weakly* closed, which by Corollary 2.7.12 of the Krein-Šmulian theorem on weakly* closed convex sets implies that $T^*(Y^*)$ is itself weakly* closed. This shows that (b) $\Rightarrow$ (c) and finishes the proof of the theorem. ■

3.1.22 Theorem. *Suppose that X and Y are Banach spaces and that $T \in B(X,Y)$.*

(a) *The operator T maps X onto Y if and only if T^* is an isomorphism from Y^* onto a subspace of X^*.*

(b) *The operator T^* maps Y^* onto X^* if and only if T is an isomorphism from X onto a subspace of Y.*

PROOF. The operator T maps X onto Y if and only if $T(X)$ is both closed and dense in Y, which by Theorems 3.1.17 and 3.1.21 is equivalent to T^* being one-to-one and having closed range, which by Corollary 1.6.6 is equivalent to T^* being an isomorphism from Y^* onto a subspace of X^*. The operator T^* maps Y^* onto X^* if and only if $T^*(Y^*)$ is both weakly* closed and weakly* dense in X^*, which is equivalent to T being one-to-one and having closed range, which is in turn equivalent to T being an isomorphism from X onto a subspace of Y. ■

See [205] for a discussion of adjoints in settings more general than that of bounded linear operators between normed spaces.

Exercises

3.1 Suppose that x^* is a bounded linear functional on a normed space X. Describe the adjoint of x^*.

3.2 Suppose that X is c_0 or ℓ_p such that $1 < p < \infty$ and that $m \in \mathbb{N}$. Let T_m and T_{-m} be the right-shift and left-shift operators defined by the formulas

$$T_m((\alpha_n)) = (0, \dots, 0, \alpha_1, \alpha_2, \dots)$$

and

$$T_{-m}((\alpha_n)) = (\alpha_{m+1}, \alpha_{m+2}, \dots),$$

where the number of leading zeros in the formula for T_m is m. It is easy to check that $T_m, T_{-m} \in B(X)$. Describe T_m^* and T_{-m}^*.

3.3 Show that there is no continuous map from $B(\mathbb{F}, c_0)$ onto $B(c_0^*, \mathbb{F}^*)$.

3.4 Suppose that X and Y are normed spaces.

(a) Prove that if $X \neq \{0\}$ and Y is not reflexive, then some member of $B(Y^*, X^*)$ is not weak*-to-weak* continuous.

(b) Prove that the isometric isomorphism $T \mapsto T^*$ maps $B(X, Y)$ onto $B(Y^*, X^*)$ if and only if either $X = \{0\}$ or Y is reflexive.

3.5 Prove that every weakly* closed subspace of the dual space of a normed space is the range of the adjoint of some bounded linear operator. That is, prove that if X is a normed space and N is a weakly* closed subspace of X^*, then there is a normed space Y and a T in $B(X, Y)$ such that $T^*(Y^*) = N$.

3.6 Prove that every weakly* closed subspace of the dual space of a normed space is the kernel of the adjoint of some bounded linear operator. That is, prove that if X is a normed space and N is a weakly* closed subspace of X^*, then there is a normed space Y and a T in $B(Y, X)$ such that $\ker(T^*) = N$. Conclude that a subspace of the dual space of a normed space is weakly* closed if and only if it is the kernel of the adjoint of some bounded linear operator.

3.7 (a) Suppose that X and Y are Banach spaces, that $T \in B(X, Y)$, and that T^* is an isomorphism from Y^* onto X^*. Prove that T^* is also a weak*-to-weak* homeomorphism from Y^* onto X^*.

(b) Give an example to show that the conclusion of (a) can fail if X and Y are only assumed to be normed spaces.

3.8 (a) If the normed space Y in Theorem 3.1.18 is incomplete, then the theorem really says nothing about T itself that was not already shown in Chapter 1, but instead amounts only to a statement about T^*. What is that statement?

(b) Show that Theorem 3.1.18 would not be true if its statement were amended to require Y instead of X to be complete.

3.9 Show that the conclusion of Theorem 3.1.21 does not have to hold if the completeness hypothesis is dropped for either X or Y.

3.10 Suppose that P is a property defined for Banach spaces such that

(1) if a Banach space X has property P, then so does every Banach space isomorphic to X;

(2) if a Banach space has property P, then so does every closed subspace of the space; and

(3) a Banach space has property P if and only if its dual space has property P.

For example, reflexivity satisfies these conditions. Prove that if X and Y are Banach spaces such that X has property P and there is a bounded linear operator from X onto Y, then Y has property P. Conclude that if a Banach space X has property P and M is a closed subspace of X, then X/M has property P.

3.11 Let X be a normed space. In this exercise, the notation Q_Y represents the natural map from Y into Y^{**} whenever Y is a normed space.

(a) Prove that Q_X^* maps X^{***} onto X^*.

(b) Prove that Q_X^{**} is an isomorphism from X^{**} onto a subspace of $X^{(4)}$.

(c) Prove that $Q_X^{**}Q_X = Q_{X^{**}}Q_X$, that is, that Q_X^{**} and $Q_{X^{**}}$ agree on $Q_X(X)$.

(d) Parts (b) and (c) might lead one to suspect that $Q_X^{**} = Q_{X^{**}}$. Prove that this is true if and only if X^* is reflexive (which is of course equivalent to X being reflexive if X is a Banach space).

3.12 Suppose that T is an isomorphism from a Banach space X onto a Banach space Y. Let Q_X and Q_Y be the natural maps from these respective Banach spaces into their second duals. Define $S\colon X^{**}/Q_X(X) \to Y^{**}/Q_Y(Y)$ by the formula $S\big(x^{**} + Q_X(X)\big) = T^{**}x^{**} + Q_Y(Y)$. Prove that S is an isomorphism from $X^{**}/Q_X(X)$ onto $Y^{**}/Q_Y(Y)$, and that S is an isometric isomorphism if T is. Notice that it must be established that S is well-defined.

3.13 Prove that a normed space is isomorphic to the dual space of a separable Banach space if and only if it is isomorphic to a weakly* closed subspace of ℓ_1^*. That is, prove that, up to isomorphism, the normed spaces that are dual spaces of separable Banach spaces are exactly the subspaces of ℓ_∞ that are weakly* closed when ℓ_∞ is identified in the usual way with ℓ_1^*.

3.2 Projections and Complemented Subspaces

The following two definitions have already been encountered in Section 1.8 for the special case in which X is a normed space.

3.2.1 Definition. Suppose that $M_1, \dots, M_n$ are closed subspaces of a topological vector space X such that $\sum_k M_k = X$ and $M_j \cap \sum_{k\neq j} M_k = \{0\}$ when $j = 1, \dots, n$. Then X is the *internal direct sum* of $M_1, \dots, M_n$.

In other words, a TVS X is the internal direct sum of its subspaces $M_1, \dots, M_n$ when each of these subspaces is closed and X is their algebraic internal direct sum. By Proposition 1.8.7, this is equivalent to requiring that each M_j be closed and that for each x in X there be unique elements $m_1(x), \dots, m_n(x)$ of $M_1, \dots, M_n$ respectively such that $x = \sum_k m_k(x)$.

3.2.2 Definition. A subspace M of a TVS X is *complemented* in X if it is closed in X and there is a closed subspace N of X such that X is the internal direct sum of M and N, in which case it is said that the two subspaces are *complementary* or that N is *complementary* to M.

By analogy with the relationship between algebraic internal direct sums and the internal direct sums of Definition 3.2.1, a subspace M of a vector space X is said to be *algebraically complemented* in X if there is a subspace N of X such that X is the algebraic internal direct sum of M and N. These more general algebraic notions will be useful in what follows, but it must be emphasized that *references in this book to internal direct sums or to the complementation of subspaces without the qualifier "algebraic" always imply that the subspaces involved are closed subspaces of some TVS.*

The question of which subspaces of a vector space are algebraically complemented can be settled quickly.

3.2.3 Proposition. *Every subspace of a vector space X is algebraically complemented in X.*

PROOF. Suppose that M is a subspace of X. Let B_M be a vector space basis for M, let B_N be a subset of X such that $B_M \cap B_N = \varnothing$ and $B_M \cup B_N$ is a basis for X, and let $N = \langle B_N \rangle$. It is easy to check that $X = M + N$ and $M \cap N = \{0\}$, so M is algebraically complemented in X. ■

Not every closed subspace of every TVS is complemented in the sense of Definition 3.2.2. In particular, it will be seen that some Banach spaces have uncomplemented closed subspaces.

Section 1.8 contained a very brief glance at complemented subspaces of normed spaces; see the discussion that follows Definition 1.8.14. The purpose of this section is to take a closer look at complemented subspaces and their relationship to linear operators of the following type.

3.2.4 Definition. Suppose that X is a vector space. A linear operator $P\colon X \to X$ is a *projection* in X if $P(Px) = Px$ for each x in X, that is, if $P^2 = P$.

Trivial examples of projections are given by the identity operator and zero operator on a vector space. The more interesting projections lie between these two extremes. In particular, the following one will provide a useful counterexample later in this section.

3.2.5 Example. Let X be the vector space of finitely nonzero sequences with the ℓ_∞ norm, let (β_n) be a sequence of scalars, and let P be the linear operator from X into X defined by the formula

$$P(\alpha_n) = (\alpha_1 + \beta_1\alpha_2, 0, \alpha_3 + \beta_2\alpha_4, 0, \alpha_5 + \beta_3\alpha_6, 0, \dots).$$

Then P is a projection. Also,

$$P(X) = \{\, (\alpha_n) : (\alpha_n) \in X,\ \alpha_{2n} = 0 \text{ for each } n \text{ in } \mathbb{N} \,\}$$

and

$$\ker(P) = \{\, (\alpha_n) : (\alpha_n) \in X,\ \alpha_{2n-1} + \beta_n\alpha_{2n} = 0 \text{ for each } n \text{ in } \mathbb{N} \,\},$$

from which it readily follows that the range and kernel of P are both closed subspaces of X and that X is the internal direct sum of these two subspaces; see Exercise 3.14. It is easy to check that P is bounded if and only if (β_n) is bounded, and that $\|P\| = 1 + \|(\beta_n)\|_\infty$ whenever (β_n) is bounded.

Notice that all of this is just as valid if X is all of ℓ_∞ provided that (β_n) is required to be bounded to assure that P takes its values in ℓ_∞, in which case P must be a bounded projection.

If $t \geq 1$, then it is possible to arrange for the projection P of the preceding example to be bounded with norm t by selecting the sequence (β_n) properly; one obvious choice is $(t-1, t-1, \dots)$. However, it is not possible for P to be bounded with norm strictly between 0 and 1. This is actually a property of all bounded projections on a normed space, since a projection always acts as the identity operator on its range.

The next several results contain some of the basic properties of projections.

3.2.6 Theorem. *Suppose that X is a vector space and that T is a linear operator from X into X. Then T is a projection if and only if the algebraic adjoint $T^\#$ of T is a projection. If X is a normed space and T is bounded, then T is a projection if and only if T^* is a projection.*

PROOF. Suppose first that $T^\#$ is a projection. For each x in X and each $x^\#$ in $X^\#$,

$$\langle T(Tx), x^\# \rangle = \langle x, T^\#(T^\# x^\#) \rangle = \langle x, T^\# x^\# \rangle = \langle Tx, x^\# \rangle,$$

which implies that $T(Tx) = Tx$ since the collection of all linear functionals on a vector space is always a separating family for that vector space. Thus, the operator T is a projection. The remaining claims in the theorem are proved by similar arguments. ■

As usual, the symbol I in the next several results represents the identity operator on the vector space in question.

3.2.7 Proposition. *Suppose that X is a vector space and that T is a linear operator from X into X. Then T is a projection if and only if $I - T$ is a projection. If X is a TVS, then T is a continuous projection if and only if $I - T$ is a continuous projection.*

PROOF. Suppose first that T is a projection. For each x in X,

$$(I - T)^2(x) = x - 2Tx + T^2x = x - Tx = (I - T)(x),$$

so $I - T$ is a projection. Conversely, if $I - T$ is a projection, then so is T since $T = I - (I - T)$. The result for TVSs now follows immediately. ■

3.2.8 Proposition. *If P is a projection in a vector space X, then* $\ker(P) = (I-P)(X)$ *and* $P(X) = \ker(I-P)$.

PROOF. If $x \in \ker(P)$, then $(I-P)(x) = x$, so $\ker(P) \subseteq (I-P)(X)$. The reverse inclusion follows from the fact that $P\big((I-P)(X)\big) = \{0\}$, which proves that $\ker(P) = (I-P)(X)$. The rest of the proposition follows by applying what has already been proved to the projection $I-P$. ■

3.2.9 Corollary. *If P is a projection in a vector space X, then $P(X) = \{\, x : x \in X,\ Px = x \,\}$.*

Since the kernel of a continuous linear operator from a Hausdorff TVS into itself must be closed, Proposition 3.2.8 has the following additional corollary.

3.2.10 Corollary. *Every continuous projection in a Hausdorff TVS has closed range.*

3.2.11 Theorem. *Suppose that X is a vector space. If P is a projection in X, then X is the algebraic internal direct sum of the range and kernel of P. Conversely, if X is the algebraic internal direct sum of its subspaces M and N, then there is a unique projection in X having range M and kernel N.*

PROOF. If P is a projection in X, then it follows from Proposition 3.2.8 that $X = P(X) + (I-P)(X) = P(X) + \ker(P)$ and $P(X) \cap \ker(P) = \ker(I-P) \cap \ker(P) = \{0\}$, so X is the algebraic internal direct sum of $P(X)$ and $\ker(P)$. Conversely, suppose that X is the algebraic internal direct sum of its subspaces M and N. By Proposition 1.8.7, every element x of X can be represented in a unique way as a sum $m(x) + n(x)$ such that $m(x) \in M$ and $n(x) \in N$. It is clear that the map $x \mapsto m(x)$ is a projection in X with range M and kernel N. Furthermore, if P_0 is any projection in X having range M and kernel N, then $P_0(x) = P_0\big(m(x)+n(x)\big) = P_0\big(m(x)\big) = m(x)$ whenever $x \in X$, which proves the uniqueness assertion. ■

3.2.12 Corollary. *If P is a continuous projection in a Hausdorff TVS X, then X is the internal direct sum of the range and kernel of P.*

Suppose that x is an element of a vector space X that is the algebraic internal direct sum of its subspaces M and N, that $P_{M,N}$ is the unique projection in X with range M and kernel N, that $P_{N,M}$ is the unique projection in X with range N and kernel M, and that m and n are the unique elements of M and N respectively such that $x = m + n$. Then $P_{M,N}(x) = m$ and $P_{N,M}(x) = n$, so $I = P_{M,N} + P_{N,M}$. Conversely, if there are projections P_1 and P_2 in X such that $I = P_1 + P_2$, then $P_2 = I - P_1$, so it follows from Theorem 3.2.11 and Proposition 3.2.8 that X is the algebraic internal direct sum of the ranges (or kernels) of P_1 and P_2. Thus, a decomposition of X as a sum of two algebraically complementary subspaces corresponds to a decomposition of I as a sum of two projections; that is, to a decomposition of I as a sum of two operators that are "smaller identities" in the sense that each acts as the identity operator on its range.

Corollary 3.2.12 suggests the following question: If M and N are complementary subspaces of a Hausdorff TVS X, must the projection with range M and kernel N be continuous? The answer does not have to be yes, even when X is a normed space.

3.2.13 Example. Suppose that X is the vector space of finitely nonzero sequences with the ℓ_∞ norm and that (β_n) is an unbounded sequence of scalars. Let P be the projection in X constructed in Example 3.2.5. Then X is the internal direct sum of the range and kernel of P even though P is not bounded.

The normed space X of the preceding example is not complete. As it turns out, that is a crucial aspect of the example.

3.2.14 Theorem. *If M and N are complementary subspaces of a Banach space X, then the projection in X with range M and kernel N is bounded.*

PROOF. Let P be the projection in question. Suppose that a sequence (x_n) in X converges to some x and that (Px_n) converges to some y. Then $(I-P)(x_n) \to x-y$. It follows that $y \in M$ and $x-y \in N$, so $y = Py = Px$. By the closed graph theorem, the operator P is bounded. ■

3.2.15 Corollary. *A subspace of a Banach space is complemented if and only if it is the range of a bounded projection in the space.*

3.2.16 Corollary. *If M and N are complementary subspaces of a Banach space X, then $M \cong X/N$.*

PROOF. Just apply the first isomorphism theorem for Banach spaces to the bounded projection in X with range M and kernel N. ■

Though the normed space version of the Hahn-Banach extension theorem guarantees that every bounded linear functional on a subspace M of a normed space X has a bounded linear extension to all of X, it does not address the more general problem of finding conditions on M, X, and normed spaces Y besides $\mathbb{F}$ that guarantee that all bounded linear operators from M into Y have bounded linear extensions to X. The next theorem, which is a straightforward consequence of Corollary 3.2.15, provides a partial solution to this problem.

3.2.17 Theorem. *Suppose that M is a subspace of a Banach space X. Then the following are equivalent.*

(a) *For every Banach space Y and every bounded linear operator T from M into Y, there is a bounded linear operator T_X from X into Y that agrees with T on M.*

(b) *The closure of M is complemented in X.*

PROOF. Suppose first that (a) holds. Let T be the identity operator on M viewed as a bounded linear operator from M into $\overline{M}$ and let T_X be a bounded linear operator from X into $\overline{M}$ that agrees with T on M. It follows from the continuity of T_X that T_X agrees with the identity operator of $\overline{M}$ on $\overline{M}$, which implies that T_X is a bounded projection in X with range $\overline{M}$. Therefore $\overline{M}$ is complemented in X by Corollary 3.2.15, which shows that (a) $\Rightarrow$ (b).

Suppose conversely that $\overline{M}$ is complemented in X. By Corollary 3.2.15, some bounded projection P in X has range $\overline{M}$. If Y is a Banach space and T is a bounded linear operator from M into Y, then Theorem 1.9.1 guarantees the existence of a bounded linear operator T_1 from $\overline{M}$ into Y that agrees with T on M, and it is then clear that T_1P is a bounded linear operator from X into Y that agrees with T on M. This proves that (b) $\Rightarrow$ (a). ■

The idea used to prove that (b) $\Rightarrow$ (a) in the preceding theorem can also be used to prove results about the extensibility of more general continuous functions. See Exercise 3.22.

The closed subspaces $\{0\}$ and X of a Hausdorff TVS X are certainly complemented, since X is the internal direct sum of the two. The following theorem says that closed subspaces of X that differ by only a finite number of dimensions from X or, if X is locally convex, from $\{0\}$, are also complemented.

Recall that if M is a subspace of a vector space X, then the *codimension* of M in X is the dimension of the quotient vector space X/M.

3.2.18 Theorem. *Suppose that M is a subspace of a Hausdorff TVS X. Then M is complemented in X if either of the following two statements holds.*

(a) *The subspace M is closed and finite-codimensional.*

(b) *The subspace M is finite-dimensional and X is locally convex.*

PROOF. It may be assumed that M is neither X nor $\{0\}$. Suppose first that M is closed and finite-codimensional. Let $x_1 + M, \dots, x_n + M$ be a basis for X/M and let $N = \langle\{x_1, \dots, x_n\}\rangle$. By Corollary 2.2.32, the finite-dimensional subspace N of X is closed. It is easy to check that $M + N = X$ and $M \cap N = \{0\}$, so M is complemented in X.

Now suppose instead that M is finite-dimensional and X is locally convex. Then M is closed by Corollary 2.2.32. Let $m_1, \ldots, m_n$ be a basis for M and let $m_j^*(\alpha_1 m_1 + \cdots + \alpha_n m_n) = \alpha_j$ whenever $j \in \{1, \ldots, n\}$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. Then each m_j^* is a linear functional on M that is continuous by Corollary 2.2.33 since M is a finite-dimensional Hausdorff TVS with respect to the topology it inherits from X. Since X is an LCS, Corollary 2.2.21 assures that each m_j^* has a continuous linear extension x_j^* to X. It follows easily that the formula $Px = (x_1^* x)m_1 + \cdots + (x_n^* x)m_n$ defines a continuous projection from X onto M, which by Corollary 3.2.12 implies that M is complemented in X. ∎

The stipulation in part (b) of the preceding theorem that X be locally convex cannot in general be omitted. See Exercise 3.20.

Not every closed subspace of every Banach space is complemented. In fact, J. Lindenstrauss and L. Tzafriri showed in 1971 [155] that the only Banach spaces having all of their closed subspaces complemented are those isomorphic to Hilbert spaces. One specific example of an uncomplemented closed subspace of a Banach space is the subspace c_0 of ℓ_∞. This was first shown in 1940 by R. S. Phillips [189]. The proof to be given here is by R. J. Whitley [244] and is based on an idea due to M. Nakamura and S. Kakutani [174] as well as A. Pełczyński and V. N. Sudakov [180]. The following lemma needed for the proof can be found in [214, p. 77]. The lemma's extremely short proof is from Whitley's paper and is credited by Whitley to Arthur Kruse.

3.2.19 Lemma. *Suppose that A is a countably infinite set. Then there is a family $\{ S_\alpha : \alpha \in I \}$ of subsets of A such that*

(1) *for each α in I, the set S_α is infinite;*

(2) *if $\alpha, \beta \in I$ and $\alpha \neq \beta$, then $S_\alpha \cap S_\beta$ is finite; and*

(3) *the index set I is uncountable.*

PROOF. It may be assumed that A is the set of rational numbers in $(0, 1)$. Let I be the irrationals in $(0, 1)$. For each α in I, let S_α be a set whose members come from a sequence of rationals in $(0, 1)$ converging to α. Then A, I, and $\{ S_\alpha : \alpha \in I \}$ do all that is required of them. ∎

3.2.20 Theorem. (R. S. Phillips, 1940 [189]). *The space c_0 is an uncomplemented closed subspace of ℓ_∞.*

PROOF. In this proof, a Banach space X will be said to have property P if X^* has a countable subset that is a separating family for X. It is clear that if a Banach space X has property P, then so does every closed subspace of X and every Banach space isomorphic to X. It is also clear that ℓ_∞ has property P; one countable separating family for ℓ_∞ in ℓ_∞^* is the collection $\{ e_n^* : n \in \mathbb{N} \}$ such that each e_n^* maps each member of ℓ_∞ to its n^{th} term.

Suppose that some closed subspace N of ℓ_∞ is complementary to the closed subspace c_0. Then $\ell_\infty/c_0 \cong N$, so ℓ_∞/c_0 has property P. A contradiction will be obtained and the theorem will be proved once it is shown that ℓ_∞/c_0 cannot have property P.

Let $A = \mathbb{N}$ and let I and $\{ S_\alpha : \alpha \in I \}$ be as in the preceding lemma. For each α in I, let x_α be the member of ℓ_∞ that, when viewed as a function from $\mathbb{N}$ into $\mathbb{F}$, is the indicator function of S_α; that is, let the n^{th} term of x_α be 1 whenever $n \in S_\alpha$ and let all other terms of x_α be 0. Notice that $x_\alpha + c_0 \neq x_\beta + c_0$ when $\alpha \neq \beta$. Suppose that $y^* \in (\ell_\infty/c_0)^*$, that $p \in \mathbb{N}$, and that $\alpha_1, \dots, \alpha_q$ are distinct members of I such that $|y^*(x_{\alpha_j} + c_0)| \geq p^{-1}$ when $j = 1, \dots, q$. Let $\gamma_1, \dots, \gamma_q$ be scalars of absolute value 1 such that $\gamma_j y^*(x_{\alpha_j} + c_0) = |y^*(x_{\alpha_j} + c_0)|$ for each j. Properties (1) and (2) listed in the lemma assure that infinitely many terms of the member $\sum_{j=1}^q \gamma_j x_{\alpha_j}$ of ℓ_∞ have absolute value 1 and that only finitely many have absolute value more than 1, from which it follows that

$$\left\| \left(\sum_{j=1}^q \gamma_j x_{\alpha_j} \right) + c_0 \right\| = d\left(\sum_{j=1}^q \gamma_j x_{\alpha_j}, c_0 \right) = 1.$$

Therefore

$$\|y^*\| \geq \left| y^*\left(\left(\sum_{j=1}^q \gamma_j x_{\alpha_j} \right) + c_0 \right) \right| = \sum_{j=1}^q |y^*(x_{\alpha_j} + c_0)| \geq p^{-1} q,$$

and so $q \leq p\|y^*\|$. Consequently, there are only finitely many index elements α such that $|y^*(x_\alpha + c_0)| \geq p^{-1}$. Since p was an arbitrary positive integer, it follows that there are only countably many index elements α such that $y^*(x_\alpha + c_0) \neq 0$.

Now suppose that C is a countable subset of $(\ell_\infty/c_0)^*$. It follows that there are only countably many members α of I such that $z^*(x_\alpha + c_0) \neq 0$ for some z^* in C. Since I is uncountable, there must be two different members α_1, α_2 of I such that $z^*(x_{\alpha_1} + c_0) = z^*(x_{\alpha_2} + c_0) = 0$ for each z^* in C, which shows that C is not a separating family for ℓ_∞/c_0. The space ℓ_∞/c_0 therefore lacks property P, which is the desired contradiction. ■

3.2.21 Corollary. *No bounded linear operator from ℓ_∞ onto c_0 maps each element of c_0 to itself.*

3.2.22 Corollary. *Let Q be the natural map from c_0 into c_0^{**}. Then $Q(c_0)$ is not complemented in c_0^{**}.*

PROOF. Suppose $Q(c_0)$ were complemented in c_0^{**}. Then there would be a bounded projection P in c_0^{**} with range $Q(c_0)$. Let $S\colon \ell_\infty \to \ell_1^*$ and $T\colon \ell_1 \to c_0^*$ be the standard identifying isometric isomorphisms, and let $P_0 = S^{-1}T^*P(T^*)^{-1}S$. Routine verifications show that P_0 is a projection in ℓ_∞ with range c_0—a contradiction. ■

The preceding corollary might lead one to ask if it is ever possible for the natural image of a nonreflexive Banach space in it second dual to be complemented. It is quite possible, as the next result shows.

3.2.23 Proposition. *Suppose that X is a normed space that is isomorphic to the dual space of a normed space and that Q_X is the natural map from X into X^{**}. Then $Q_X(X)$ is complemented in X^{**}.*

PROOF. Since the dual space of a normed space is a Banach space that is isometrically isomorphic to the dual space of the completion of its predual, the space X is actually a Banach space isomorphic to the dual space Y^* of a Banach space Y. Let T be an isomorphism from X onto Y^* and let Q_{Y^*} be the natural map from Y^* into Y^{***}. It is enough to show that there is a bounded projection P in Y^{***} with range $Q_{Y^*}(Y^*)$, for then $(T^{**})^{-1}PT^{**}$ is a bounded projection in X^{**} with range $Q_X(X)$ since

$$(T^{**})^{-1}PT^{**}(T^{**})^{-1}PT^{**} = (T^{**})^{-1}PPT^{**} = (T^{**})^{-1}PT^{**}$$

and

$$(T^{**})^{-1}PT^{**}(X^{**}) = (T^{-1})^{**}P(Y^{***}) = (T^{-1})^{**}Q_{Y^*}(Y^*) = Q_X(X),$$

where the very last equality comes from Corollary 3.1.14.

Let Q_Y be the natural map from Y into Y^{**} and let $P = Q_{Y^*}Q_Y^*$. For each y in Y and each y^* in Y^*,

$$\langle y, Q_Y^*Q_{Y^*}y^*\rangle = \langle Q_Y y, Q_{Y^*}y^*\rangle = \langle y^*, Q_Y y\rangle = \langle y, y^*\rangle,$$

so $Q_Y^*Q_{Y^*}$ is the identity map on Y^*. It follows that $P^2 = P$, that is, that P is a projection in Y^{***}. Since Q_Y is an isomorphism from the Banach space Y onto a subspace of Y^{**}, its adjoint Q_Y^* maps Y^{***} onto Y^*, so $P(Y^{***}) = Q_{Y^*}Q_Y^*(Y^{***}) = Q_{Y^*}(Y^*)$. ■

3.2.24 Corollary. *The Banach space c_0 is not isomorphic to the dual space of any normed space.*

The preceding corollary improves the result of Example 2.10.11, where the fact that B_{c_0} has no extreme points was used to show that c_0 is not *isometrically* isomorphic to the dual space of any normed space.

Exercises

3.14 Provide the missing details in Example 3.2.5 by showing that the range and kernel of P are both closed subspaces of X, that X is the internal direct sum of these two subspaces, that P is bounded if and only if (β_n) is bounded, and that $\|P\| = 1 + \|(\beta_n)\|_\infty$ whenever (β_n) is bounded.

3.15 For each T in $L(\mathbb{F}^2, \mathbb{F}^2)$, let the *determinant* of T be the determinant of the matrix $m(T)$ of T with respect to the standard basis for $\mathbb{F}^2$, and let the *trace* of T be the trace of $m(T)$, that is, the sum of the elements along the main diagonal of $m(T)$. Prove that a member T of $L(\mathbb{F}^2, \mathbb{F}^2)$ is a projection if and only if it satisfies one of the following three conditions.

(a) $T = I$.

(b) $T = 0$.

(c) The determinant of T is 0 and the trace of T is 1.

3.16 (a) Prove that if M is a complemented subspace of a Banach space, then all subspaces complementary to M are isomorphic to one another.

(b) Give an example of a complemented subspace of a Banach space that has infinitely many different subspaces complementary to it.

3.17 (a) Suppose that P is a bounded projection in a normed space X. Prove that $P(X) = {}^{\perp}(\ker(P^*))$ and $P^*(X^*) = (\ker(P))^{\perp}$.

(b) Suppose that M and N are complementary subspaces of a Banach space X. Prove that $M^{\perp}$ and $N^{\perp}$ are complementary in X^*.

(c) Give an example to show that the conclusion of (b) need not hold if X is only assumed to be a normed space.

3.18 Suppose that P is a bounded projection in a normed space X. Prove that $P(X)$ has finite dimension n if and only if $P^*(X^*)$ has finite dimension n. (One proof of this uses Exercise 3.17 (a) and either Exercise 1.81 or the first isomorphism theorem for vector spaces; see, for example, [105, p. 397] for this theorem.)

3.19 This exercise uses the material of optional Section 2.3. Show that Theorem 3.2.14 and Corollary 3.2.15 have natural extensions to F-spaces.

3.20 Suppose that $0 \le p < 1$. Prove that no finite-dimensional subspace of $L_p[0, 1]$ other than $\{0\}$ is complemented. Exercise 3.19 may be helpful.

3.21 This exercise improves the conclusion of Exercise 1.141. Suppose that X is a Banach space and that there is a bounded linear operator from X onto ℓ_1. Prove that X has a complemented subspace isomorphic to ℓ_1.

3.22 Suppose that M is a subspace of a Banach space X and that $\overline{M}$ is complemented in X.

(a) Prove that if M is closed, then the following holds: For every topological space Z and every continuous function f from M into Z, there is a continuous function f_X from X into Z that agrees with f on M.

(b) Show by example that the conclusion in (a) does not necessarily hold if M is not required to be closed, even when Z is assumed to be a normed space and f is assumed to be linear.

3.23 Suppose that M is a closed subspace of a Banach space. It follows from Exercise 3.17 (b) that if M is complemented, then so is $M^{\perp}$. The main goal of this exercise is to prove that the converse is not in general true.

(a) Prove that if X is a Banach space and Q_X and Q_{X^*} are the natural maps from X and X^* into X^{**} and X^{***} respectively, then $Q_{X^*}(X^*)$ and $\big(Q_X(X)\big)^{\perp}$ are complementary subspaces of X^{***}.

(b) Find a closed subspace M of a Banach space such that $M^{\perp}$ is complemented but M is not.

3.24 No result in this section says anything about complemented subspaces of TVSs that are not Hausdorff. There is a good reason for this. Find it.

3.3 Banach Algebras and Spectra

The main purpose of this section is to obtain a few facts about the spectrum of a bounded linear operator from a Banach space into itself. Most of the discussion of the spectrum will take place within the more general context of Banach algebras, since it is no more difficult to do so and since it provides an opportunity to take a brief look at these objects.

3.3.1 Definition. Suppose that X is a set, that $+$ and $\bullet$ are binary operations from $X \times X$ into X, and that $\cdot$ is a binary operation from $\mathbb{F} \times X$ into X such that

(1) $(X, +, \cdot)$ is a vector space;

and for all x, y, z in X and every scalar α,

(2) $x \bullet (y \bullet z) = (x \bullet y) \bullet z$;

(3) $x \bullet (y + z) = (x \bullet y) + (x \bullet z)$ and $(x + y) \bullet z = (x \bullet z) + (y \bullet z)$; and

(4) $\alpha \cdot (x \bullet y) = (\alpha \cdot x) \bullet y = x \bullet (\alpha \cdot y)$.

Then $(X, +, \bullet, \cdot)$ is an *algebra*. This algebra is an *algebra with identity* if $X \neq \{0\}$ and there is a member e of X such that

(5) $e \bullet x = x \bullet e = x$

whenever $x \in X$, in which case e is the *(multiplicative) identity* of X. If $\|\cdot\|$ is a norm on the vector space $(X, +, \cdot)$ such that

(6) $\|x \bullet y\| \leq \|x\| \, \|y\|$

whenever $x, y \in X$, then $\big(X, +, \bullet, \cdot, \|\cdot\|\big)$ is a *normed algebra,* and is a *Banach algebra* if the norm is a Banach norm.

With all notation as in the preceding definition, the products $\alpha \cdot x$ and $x \bullet y$ are usually abbreviated to αx and xy respectively, and it is usually said that X is an algebra (or normed algebra or Banach algebra if there is a norm involved) rather than using the more formal notations $(X, +, \bullet, \cdot)$ and $\big(X, +, \bullet, \cdot, \|\cdot\|\big)$.

The reason for requiring an algebra with identity to contain more than 0 is to make it impossible for 0 to be a multiplicative identity for the algebra. See the proof of Proposition 3.3.10 (b).

3.3.2 Example. The set $\mathbb{F}$ with its usual vector space operations, its usual multiplication, and the norm given by the absolute value function is a Banach algebra with identity 1.

3.3.3 Example. Suppose that K is a compact Hausdorff space. Define the product of two members f and g of $C(K)$ to be the usual pointwise product. Then $C(K)$ is a Banach algebra when given this multiplication, the usual vector space operations, and the usual norm. As long as K is nonempty, the algebra $C(K)$ has an identity, namely, the function taking on the constant value 1 on K.

3.3.4 Example. Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω. As with $C(K)$, let the product of two members f and g of $L_\infty(\Omega, \Sigma, \mu)$ be the pointwise product. Then $L_\infty(\Omega, \Sigma, \mu)$ with this multiplication, the usual vector space operations, and the usual norm is a Banach algebra. As long as $\mu(\Omega) > 0$, the function having constant value 1 on Ω is an identity for $L_\infty(\Omega, \Sigma, \mu)$.

3.3.5 Example. As a special case of the preceding example, the space ℓ_∞ is a Banach algebra with identity $(1, 1, 1, \dots)$. Notice that the multiplication of elements of ℓ_∞ is done termwise.

3.3.6 Example. With ℓ_∞ treated as a Banach algebra as in the preceding example, the space c_0 is a closed subalgebra of ℓ_∞ and therefore is itself a Banach algebra. Notice that c_0 has no multiplicative identity.

The following example is the most important one for the purposes of this book.

3.3.7 Example. Suppose that X is a normed space. As has been done throughout this book, let the product ST of two members S and T of $B(X)$ be given by composition, that is, by the formula $ST(x) = S\big(T(x)\big)$. Then $B(X)$ with this multiplication, the usual vector space operations, and the operator norm is a normed algebra. If X is a Banach space, then $B(X)$ is a Banach algebra. As long as $X \neq \{0\}$, the identity operator on X is a multiplicative identity for $B(X)$.

The multiplication operations defined in the above examples are the standard ones for their spaces and are the ones that will be assumed when these spaces are treated as algebras in this book.

3.3.8 Definition. Suppose that X is an algebra with identity e. A member x of X is *invertible* if there is a y in X such that $xy = yx = e$, in which case y is the *(multiplicative) inverse* of x and is denoted by x^{-1}.

3.3.9 Definition. Suppose that x is an element of an algebra X. Then x^n is defined inductively for each positive integer n by letting $x^1 = x$ and $x^n = x^{n-1}x$ when $n \geq 2$. Now suppose that X has identity e. Then x^0 is defined to be e. If x is invertible and $n \in \mathbb{N}$, then x^{-n} is defined to be $(x^{-1})^n$.

It is implicitly assumed in the preceding two definitions, as well as in the use of the phrase "*the* (multiplicative) identity" in Definition 3.3.1, that every algebra has at most one identity and that an invertible element of an algebra with identity has at most one inverse. This does turn out to be the case. The next proposition contains these facts, as well as some other basic ones about algebras.

3.3.10 Proposition. *Suppose that X is an algebra.*

(a) *If $x \in X$ and 0 is the zero element of X, then $0x = x0 = 0$.*

(b) *The element 0 of X is not a multiplicative identity for X.*

(c) *The algebra X has at most one multiplicative identity.*

Now suppose that X is an algebra with identity e.

(d) *The element 0 of X is not invertible.*

(e) *Each element of X has at most one inverse.*

(f) *If x, y, and z are elements of X such that $yx = xz = e$, then $y = z$ and x is invertible with inverse y. That is, every element of X that is both left-invertible and right-invertible is invertible, and each of its left inverses and right inverses equals its inverse.*

(g) *If x and y are invertible elements of X and α is a nonzero scalar, then xy, αx, and x^{-1} are invertible and have respective inverses $y^{-1}x^{-1}$, $\alpha^{-1}x^{-1}$, and x.*

(h) *If x is an invertible element of X and $n \in \mathbb{N}$, then x^n is invertible, and $(x^n)^{-1} = (x^{-1})^n = x^{-n}$.*

Suppose next that X is a normed algebra, possibly without identity.

(i) *If $x \in X$ and $n \in \mathbb{N}$, then $\|x^n\| \leq \|x\|^n$.*

Finally, suppose that X is a normed algebra with identity e.

(j) $\|e\| \geq 1$.

(k) *If x is an invertible element of X, then $\|x^{-1}\| \geq \|x\|^{-1}$.*

PROOF. If $x \in X$ and 0 is the zero element of X, then $0x = (0+0)x = 0x + 0x$, so $0 = 0x$, and similarly $0 = x0$. This gives (a), from which (b) follows immediately once it is noted that, by definition, an algebra with identity must have a nonzero element. If e_1 and e_2 are both identities for X, then $e_1 = e_1 e_2 = e_2$, proving (c).

Suppose that X has identity e. It follows immediately from (a) and (b) that 0 is not invertible, giving (d). If an element x of X has inverses y and z, then $y = ye = y(xz) = (yx)z = ez = z$, proving (e). The same string of equalities proves (f). Parts (g) and (h) are easily verified.

Now suppose that X is a normed algebra, possibly having no identity, and that $x \in X$ and n is a positive integer greater than 1. Then $\|x^1\| = \|x\|^1$ and $\|x^n\| \leq \|x^{n-1}\| \, \|x\|$, which together with an obvious induction argument yields (i).

Finally, suppose that X is a normed algebra having identity e. Then $\|e\| = \|e^2\| \leq \|e\|^2$, which together with the fact that $e \neq 0$ proves (j). If x is an invertible element of X, then $1 \leq \|e\| = \|x^{-1}x\| \leq \|x^{-1}\| \, \|x\|$, so $\|x^{-1}\| \geq \|x\|^{-1}$, which finishes the proof of (k) and of the proposition. ■

Many of the familiar laws of exponents from basic arithmetic also hold for algebras, and are easy consequences of Definition 3.3.9 and parts (g) and (h) of Proposition 3.3.10 (but beware of assuming that $(xy)^n = x^n y^n$ unless it is known that $xy = yx$).

It often happens that an identity for a normed algebra has norm 1, and the examples given in this section might lead one to suspect that this must always be the case. However, part (j) of Proposition 3.3.10 cannot in general be sharpened, even for Banach algebras. See Exercises 3.27 and 3.28.

Addition of vectors and multiplication of vectors by scalars are continuous operations for a normed algebra because of the continuity of these operations for normed spaces. It turns out that multiplication of vectors by vectors is also continuous. The proof is essentially the same as the familiar proof of the continuity of multiplication in the scalar field.

3.3.11 Proposition. *The multiplication of members of a normed algebra X is a continuous operation from $X \times X$ into X.*

PROOF. Suppose that (x_n) and (y_n) are sequences in X for which there are elements x and y of X such that $x_n \to x$ and $y_n \to y$. Since the sequence (x_n) is bounded,

$$\begin{aligned} \|x_n y_n - xy\| &\leq \|x_n y_n - x_n y\| + \|x_n y - xy\| \\ &\leq \|x_n\| \, \|y_n - y\| + \|y\| \, \|x_n - x\| \to 0 \end{aligned}$$

as $n \to \infty$, so $x_n y_n \to xy$ as required. ■

The continuity of the inversion map $x \mapsto x^{-1}$ on its domain in a Banach algebra is most easily treated after the introduction of resolvent functions and will be taken up then.

One immediate consequence of the preceding proposition is that multiplication in a normed algebra distributes through infinite sums; that is, that if x is a member of a normed algebra and $\sum_n x_n$ is a convergent series in the algebra, then $\sum_n xx_n$ and $\sum_n x_n x$ both converge, and $x(\sum_n x_n) = \sum_n xx_n$ and $(\sum_n x_n)x = \sum_n x_n x$. Here is an application of this.

Suppose that X is a Banach algebra with identity e, that $x \in X$, and that α is a scalar such that $|\alpha| > \|x\|$. Then

$$\begin{aligned}\sum_{n=0}^{\infty} \|\alpha^{-(n+1)}x^n\| &= |\alpha|^{-1}\left(\|e\| + \sum_{n=1}^{\infty} \|(\alpha^{-1}x)^n\|\right) \\ &\le |\alpha|^{-1}\left(\|e\| + \sum_{n=1}^{\infty} \|\alpha^{-1}x\|^n\right).\end{aligned}$$

This last series is a convergent geometric series, so $\sum_{n=0}^{\infty} \alpha^{-(n+1)}x^n$ is an absolutely convergent series that is therefore convergent. This, together with the formula for the sum of a convergent geometric series of scalars, suggests that $\alpha e - x$ might be invertible with inverse $\sum_{n=0}^{\infty} \alpha^{-(n+1)}x^n$, which is in fact the case since

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

and similarly $(\sum_{n=0}^{\infty} \alpha^{-(n+1)}x^n)(\alpha e - x) = e$. This motivates the next definition and proves the theorem that follows it.

3.3.12 Definition. Suppose that x is an element of an algebra X with identity e. Then the *resolvent set* $\rho(x)$ of x is the set of all scalars α such that $\alpha e - x$ is invertible. The *spectrum* $\sigma(x)$ of x is the complement in $\mathbb{F}$ of $\rho(x)$. The *resolvent function* or *resolvent* R_x of x is the function from $\rho(x)$ into X defined by the formula $R_x(\alpha) = (\alpha e - x)^{-1}$.

3.3.13 Theorem. *Suppose that x is an element of a Banach algebra with identity and that α is a scalar such that $|\alpha| > \|x\|$. Then $\alpha \in \rho(x)$ and $R_x(\alpha) = \sum_{n=0}^{\infty} \alpha^{-(n+1)}x^n$.*

3.3.14 Corollary. *If x is an element of a Banach algebra with identity and $\alpha \in \sigma(x)$, then $|\alpha| \le \|x\|$.*

3.3.15 Corollary. *If x is an element of a Banach algebra with identity e and $\|x - e\| < 1$, then x is invertible and $x^{-1} = \sum_{n=0}^{\infty}(e - x)^n$.*

PROOF. Since $x = 1e - (e - x)$ and $|1| > \|e - x\|$, Theorem 3.3.13 assures the invertibility of x and yields the series expansion for x^{-1} from the series expansion for the resolvent function. ∎

All the ingredients are now at hand for the examination of the continuity of inversion promised earlier in this section.

3.3.16 Proposition. *Suppose that X is a Banach algebra with identity e. Let $G(X)$ be the set of invertible elements of X. Then $G(X)$ is an open subset of X that is a group with identity e under the restriction of the multiplication operation of X to $G \times G$. Furthermore, the map $x \mapsto x^{-1}$ is a topological homeomorphism of $G(X)$ onto itself.*

PROOF. It is clear that $G(X)$ has the claimed group structure. By parts (e) and (g) of Proposition 3.3.10, the map $x \mapsto x^{-1}$ is one-to-one from $G(X)$ onto $G(X)$ and is its own inverse, so the proposition will be proved once it is shown that the map is continuous and that $G(X)$ is open. Suppose that $x \in G(X)$ and that $\epsilon > 0$. It follows from the continuity of multiplication of members of X that there is a positive δ such that $\|x^{-1}y\| < \epsilon/(\|x^{-1}\| + \epsilon)$ whenever $y \in X$ and $\|y\| < \delta$. Fix a z in X such that $\|z\| < \delta$. It is enough to prove that $x + z \in G(X)$ and $\|(x+z)^{-1} - x^{-1}\| < \epsilon$.

Since $\|-x^{-1}z\| < 1$, Theorem 3.3.13 assures that $1 \in \rho(-x^{-1}z)$ and therefore that $e + x^{-1}z$ is invertible and has inverse $\sum_{n=0}^{\infty}(-x^{-1}z)^n$. It follows that $x + z$ is invertible because $x + z = x(e + x^{-1}z)$. Furthermore,

$$\begin{aligned}
\|(x+z)^{-1} - x^{-1}\| &= \|(e + x^{-1}z)^{-1}x^{-1} - x^{-1}\| \\
&\le \|x^{-1}\|\,\|(e + x^{-1}z)^{-1} - e\| \\
&= \|x^{-1}\| \left\| \sum_{n=0}^{\infty} (-x^{-1}z)^n - e \right\| \\
&= \|x^{-1}\| \left\| \sum_{n=1}^{\infty} (-x^{-1}z)^n \right\| \\
&\le \|x^{-1}\| \sum_{n=1}^{\infty} \|x^{-1}z\|^n \\
&< \|x^{-1}\| \sum_{n=1}^{\infty} \left(\frac{\epsilon}{\|x^{-1}\| + \epsilon} \right)^n \\
&= \epsilon,
\end{aligned}$$

as required. ■

With all notation as in the preceding proposition, suppose that $x \in X$. Then $\rho(x) = \{\, \alpha : \alpha \in \mathbb{F},\ \alpha e - x \in G(X) \,\}$. The next result is therefore an immediate consequence of the continuity of the map $\alpha \mapsto \alpha e - x$ from $\mathbb{F}$ into X.

3.3.17 Proposition. *If x is an element of a Banach algebra with identity, then $\rho(x)$ is open.*

It follows from the preceding proposition and Corollary 3.3.14 that the spectrum of an element of a Banach algebra with identity is closed and bounded. This can be summarized as follows.

3.3.18 Theorem. *If x is an element of a Banach algebra with identity, then $\sigma(x)$ is compact.*

It is possible for the spectrum of an element of a Banach algebra with identity to be empty if the Banach algebra is real; see Exercise 3.30. As will now be shown, this never happens when the scalar field is $\mathbb{C}$. The following fact will be useful for one of the lemmas needed for the proof.

3.3.19 The Resolvent Equation. *If x is an element of an algebra with identity, then*

$$R_x(\alpha) - R_x(\beta) = (\beta - \alpha)R_x(\alpha)R_x(\beta)$$

whenever $\alpha, \beta \in \rho(x)$.

PROOF. Let e denote the identity of the algebra. If $\alpha, \beta \in \rho(x)$, then

$$\begin{aligned}(\beta - \alpha)R_x(\alpha)R_x(\beta) &= (\alpha e - x)^{-1}(\beta e - \alpha e)(\beta e - x)^{-1} \\ &= (\alpha e - x)^{-1}\big((\beta e - x) - (\alpha e - x)\big)(\beta e - x)^{-1} \\ &= (\alpha e - x)^{-1} - (\beta e - x)^{-1} \\ &= R_x(\alpha) - R_x(\beta),\end{aligned}$$

as claimed. ■

3.3.20 Definition. Suppose that f is a function from an open subset G of $\mathbb{F}$ into a normed space X, that $\alpha \in G$, and that

$$\lim_{\beta \to 0} \left(\frac{1}{\beta}\big(f(\alpha + \beta) - f(\alpha)\big) \right)$$

exists. Then f is *differentiable* at α. The above limit is the *derivative* of f at α and is denoted by $f'(\alpha)$.

3.3.21 Lemma. *If x is an element of a Banach algebra with identity e, then R_x is differentiable on $\rho(x)$, and $R_x'(\alpha) = -(\alpha e - x)^{-2}$ whenever $\alpha \in \rho(x)$.*

PROOF. Suppose that $\alpha \in \rho(x)$. Let β be a nonzero scalar such that $\alpha + \beta \in \rho(x)$. It follows from the resolvent equation that

$$\begin{aligned}\frac{1}{\beta}\big(R_x(\alpha + \beta) - R_x(\alpha)\big) &= -R_x(\alpha + \beta)R_x(\alpha) \\ &= -\big((\alpha + \beta)e - x\big)^{-1}\big(\alpha e - x\big)^{-1}.\end{aligned}$$

Letting β tend to 0 establishes the lemma. ■

3.3.22 Lemma. *Suppose that x is an element of a Banach algebra X with identity and that $x^* \in X^*$. Then $x^* R_x$ is differentiable on $\rho(x)$.*

PROOF. Let e denote the identity of X. It follows from the preceding lemma and the continuity and linearity of x^* that $x^* R_x$ is differentiable on $\rho(x)$ with derivative $x^*\big(-(\alpha e - x)^{-2}\big)$. ■

3.3.23 Lemma. *If x is an element of a Banach algebra with identity, then $\lim_{|\alpha|\to\infty} R_x(\alpha) = 0$.*

PROOF. Let e be the identity of the Banach algebra. If α is a scalar such that $|\alpha| > \|x\|$, then it follows from Theorem 3.3.13 and the computation preceding Definition 3.3.12 that

$$\begin{aligned}\|R_x(\alpha)\| &= \left\|\sum_{n=0}^{\infty} \alpha^{-(n+1)} x^n\right\| \\ &\le \sum_{n=0}^{\infty} \|\alpha^{-(n+1)} x^n\| \\ &\le |\alpha|^{-1}\left(\|e\| + \sum_{n=1}^{\infty} \|\alpha^{-1}x\|^n\right) \\ &= |\alpha|^{-1}\left(\|e\| + \frac{\|x\|}{|\alpha| - \|x\|}\right).\end{aligned}$$

Letting $|\alpha|$ tend to ∞ finishes the proof. ■

3.3.24 Theorem. *If x is an element of a complex Banach algebra with identity, then $\sigma(x)$ is nonempty.*

PROOF. Let X be the Banach algebra. Suppose that $\sigma(x) = \varnothing$. For each x^* in X^*, it follows from Lemmas 3.3.22 and 3.3.23 that $x^* R_x$ is bounded and entire and so, by Liouville's theorem, constant; another application of Lemma 3.3.23 then shows that the constant value is 0. This assures that $R_x(\alpha) = 0$ for each α in $\mathbb{C}$, which implies that 0 is invertible. Since this cannot be, the supposition that $\sigma(x) = \varnothing$ must be wrong. ■

Thus, each element x of a complex Banach algebra with identity has a nonempty compact spectrum lying in the closed disc in $\mathbb{C}$ centered at 0 of radius $\|x\|$. The next portion of this section is devoted to finding the radius of the smallest closed disc in $\mathbb{C}$ centered at 0 that includes the spectrum.

3.3.25 Definition. Suppose that x is an element of a complex Banach algebra with identity. Let

$$r_\sigma(x) = \max\{\, |\alpha| : \alpha \in \sigma(x) \,\}.$$

Then $r_\sigma(x)$ is the *spectral radius* of x.

3.3.26 Lemma. *Suppose that x is an element of an algebra with identity, that $n \in \mathbb{N}$, and that $\alpha \in \sigma(x)$. Then $\alpha^n \in \sigma(x^n)$.*

PROOF. Let e be the identity of the algebra. Suppose that β is a scalar such that $\beta^n \notin \sigma(x^n)$, that is, such that $\beta^n e - x^n$ is invertible. The lemma will be proved once it is shown that $\beta e - x$ is also invertible. Since

$$\beta^n e - x^n = (\beta e - x)\Big(\sum_{j=0}^{n-1} \beta^{n-1-j} x^j\Big) = \Big(\sum_{j=0}^{n-1} \beta^{n-1-j} x^j\Big)(\beta e - x)$$

(where β^0 is to be interpreted to be 1 if $\beta = 0$), it follows that

$$\begin{aligned} e &= (\beta e - x)\Big(\sum_{j=0}^{n-1} \beta^{n-1-j} x^j\Big)(\beta^n e - x^n)^{-1} \\ &= (\beta^n e - x^n)^{-1}\Big(\sum_{j=0}^{n-1} \beta^{n-1-j} x^j\Big)(\beta e - x). \end{aligned}$$

Since $\beta e - x$ has both a right inverse and a left inverse, it is invertible. ■

In the following result, the existence of the limit in the formula is part of the conclusion. The existence of the limit does not actually require the complex normed algebra in question to be complete or to have an identity; see Exercise 3.31.

3.3.27 The Spectral Radius Formula. *If X is a complex Banach algebra with identity, then*

$$r_\sigma(x) = \lim_n \|x^n\|^{1/n}$$

whenever $x \in X$.

PROOF. Fix an x in X. Suppose first that $\beta \in \sigma(x)$. For each positive integer n, the preceding lemma assures that $\beta^n \in \sigma(x^n)$ and therefore that $|\beta^n| \le \|x^n\|$ and $|\beta| \le \|x^n\|^{1/n}$. It follows that $|\beta| \le \liminf_n \|x^n\|^{1/n}$ and thus that $r_\sigma(x) \le \liminf_n \|x^n\|^{1/n}$.

Now suppose that γ is a scalar such that $|\gamma| > r_\sigma(x)$. Let x^* be a member of X^*. Then $x^* R_x$ is analytic on $\{\, \alpha : \alpha \in \mathbb{C},\ |\alpha| > r_\sigma(x) \,\}$ by Lemma 3.3.22. Applying x^* to the series expansion for R_x from Theorem 3.3.13 shows that $x^* R_x(\alpha) = \sum_{n=0}^\infty x^*(x^n)\alpha^{-(n+1)}$ when $|\alpha| > \|x\|$. This series expansion must actually hold for $x^* R_x(\alpha)$ when $|\alpha| > r_\sigma(x)$ by the uniqueness of Laurent series expansions on annuli; see, for example, [42, p. 103]. In particular, it holds for $x^* R_x(\gamma)$, so $\sum_{n=0}^\infty x^*(x^n)\gamma^{-(n+1)}$ converges, which in turn implies that $\lim_n x^*(\gamma^{-(n+1)} x^n) = 0$. Since x^* is an arbitrary member of X^*, it follows that $\gamma^{-(n+1)} x^n \xrightarrow{w} 0$ and therefore that the sequence $(\gamma^{-(n+1)}x^n)$ is bounded. Let B be a positive upper bound for the norms of the terms of this sequence. For each positive integer n, it follows that $\|x^n\| \le |\gamma|^{n+1}B$ and thus that $\|x^n\|^{1/n} \le |\gamma|(|\gamma|B)^{1/n}$. Therefore $\limsup_n \|x^n\|^{1/n} \le |\gamma|$ since $\lim_n(|\gamma|B)^{1/n} = 1$. Since γ could have been any scalar with absolute value greater than $r_\sigma(x)$, it must be that $\limsup_n \|x^n\|^{1/n} \le r_\sigma(x)$, which when combined with the inequality involving $\liminf_n \|x^n\|^{1/n}$ derived in the first part of this proof shows that $\lim_n \|x^n\|^{1/n}$ exists and equals $r_\sigma(x)$. ■

3.3.28 Corollary. *Suppose that X is a complex Banach algebra with identity and that M is a closed subalgebra of X with an identity (not assumed to be the identity of X). Let x be an element of M and let $r_{\sigma_M}(x)$ and $r_{\sigma_X}(x)$ be the spectral radii of x with respect to the algebras M and X respectively. Then $r_{\sigma_M}(x) = r_{\sigma_X}(x)$.*

Notice that the preceding corollary says only that the spectral radii of x with respect to the two algebras have to be the same, not that the spectra themselves do. See Exercise 3.34.

This ends the portion of this section covering the general theory of Banach algebras. Those interested in learning more about this theory will find [148] to be an excellent source.

The rest of this section is devoted to a very brief look at the particular Banach algebras of the most interest in this chapter, namely, the Banach algebras $B(X)$ such that X is a *nontrivial* Banach space, that is, a Banach space that is not just $\{0\}$. As usual, the symbol I will be used throughout this discussion and the rest of this chapter to denote the identity operator on X. The identity operator on X^* will be denoted by I^*, which is consistent with the fact that this operator is the adjoint of I. Notice that I and I^* are multiplicative identities for $B(X)$ and $B(X^*)$ respectively since X is nontrivial. Notice also that a member T of $B(X)$ is invertible in the sense of Definition 3.3.8 if and only if it is a normed space isomorphism from X onto itself, which by Corollary 1.6.6 happens exactly when T is *algebraically* invertible, that is, when T is one-to-one and maps X onto itself. Incidentally, this would not in general be true if X were not assumed to be complete, since it is possible for a one-to-one bounded linear operator from an incomplete normed space onto itself not to have a bounded inverse; consider the map $(\alpha_n) \mapsto (n^{-1}\alpha_n)$ on the vector space of finitely nonzero sequences with the ℓ_∞ norm.

Suppose that $T \in B(X)$, where X is a nontrivial Banach space, and that $\alpha \in \mathbb{F}$. Since $(\alpha I - T)^* = \alpha I^* - T^*$, it follows from Theorem 3.1.18 that $\alpha I - T$ is invertible if and only if $\alpha I^* - T^*$ is invertible. If $\alpha I - T$ is invertible, then an application of Proposition 3.1.15 shows that $\big((\alpha I - T)^{-1}\big)^* = (\alpha I^* - T^*)^{-1}$. These observations can be summarized as follows.

3.3.29 Proposition. *Suppose that X is a nontrivial Banach space and that $T \in B(X)$. Then $\sigma(T) = \sigma(T^*)$. If $\alpha \in \rho(T)$, then $\big(R_T(\alpha)\big)^* = R_{T^*}(\alpha)$.*

If $T \in B(X)$, where X is a nontrivial Banach space, then a scalar α is in the spectrum of T if and only if $\alpha I - T$ is not invertible, that is, if and only if $\alpha I - T$ either is not one-to-one or fails to map X onto X. The first possibility is interesting enough to deserve special treatment.

3.3.30 Definition. Suppose that X is a nontrivial Banach space and that $T \in B(X)$. The *point spectrum* $\sigma_p(T)$ of T is the collection of all scalars α such that $\alpha I - T$ is not one-to-one. That is,

$$\sigma_p(T) = \{\, \alpha : \alpha \in \mathbb{F},\ \ker(\alpha I - T) \neq \{0\}\,\}.$$

Each member of $\sigma_p(T)$ is an *eigenvalue* of T. If $\alpha \in \sigma_p(T)$, then the closed subspace $\ker(\alpha I - T)$ of X is the *eigenspace* associated with the eigenvalue α of T, and each nonzero member of this eigenspace is an *eigenvector* of T.

When X and T are as in the preceding definition, the *continuous spectrum* $\sigma_c(T)$ of T is the collection of all scalars α such that $\alpha I - T$ is one-to-one and $(\alpha I - T)(X)$ is a proper dense subset of X, while the *residual spectrum* $\sigma_r(T)$ of T is the collection of all scalars α such that $\alpha I - T$ is one-to-one and $(\alpha I - T)(X)$ is not dense in X. These two terms will not be needed here. Notice that $\sigma(T)$ is the disjoint union of $\sigma_p(T)$, $\sigma_c(T)$, and $\sigma_r(T)$.

Eigenvalues have many other names (with corresponding names for eigenspaces and eigenvectors), including *proper values, characteristic values, characteristic roots, secular values, spectral values, latent values,* and *latent roots.* See [67, pp. 606–607] for an interesting quote by Sylvester concerning the rationale for that last name.

With X and T as in Definition 3.3.30, a scalar α is an eigenvalue of T if and only if there is a nonzero x in X such that $Tx = \alpha x$, while a nonzero member x of X is an eigenvector of T if and only if T maps x to a scalar multiple of itself. Some sources allow the zero element of X to be an eigenvector. While this does make every member of an eigenspace an eigenvector, many interesting facts about eigenvectors hold only for nonzero ones anyway. For example, see the next theorem.

Though the spectrum of a bounded linear operator from a nontrivial complex Banach space into itself must be nonempty, it is quite possible for its point spectrum to be empty. See Exercise 3.38.

The following theorem is just another way to say that if T is a bounded linear operator from a nontrivial Banach space into itself, then every sum of a nonempty finite collection of distinct eigenspaces of T is actually an internal direct sum.

3.3.31 Theorem. *Suppose that X is a nontrivial Banach space, that $T \in B(X)$, that $\alpha_1, \dots, \alpha_n$ are distinct eigenvalues of T, and that $x_1, \dots, x_n$ are eigenvectors associated with $\alpha_1, \dots, \alpha_n$ respectively. Then $x_1, \dots, x_n$ are linearly independent.*

PROOF. It may be assumed that $n > 1$. Suppose that $j \in \{2, \ldots, n\}$ and that $x_1, \ldots, x_{j-1}$ are linearly independent. Suppose moreover that there are scalars $\beta_1, \ldots, \beta_{j-1}$ such that $x_j = \beta_1 x_1 + \cdots + \beta_{j-1} x_{j-1}$. Then

$$0 = (\alpha_j I - T)(x_j) = \beta_1(\alpha_j - \alpha_1)x_1 + \cdots + \beta_{j-1}(\alpha_j - \alpha_{j-1})x_{j-1}.$$

Since $\alpha_j - \alpha_k \neq 0$ when $k = 1, \ldots, j-1$, it follows that $\beta_1 = \cdots = \beta_{j-1} = 0$. This implies that $x_j = 0$, a contradiction. It must be that $x_1, \ldots, x_j$ are linearly independent, so it follows by induction that $x_1, \ldots, x_n$ are linearly independent. ■

The linear independence of eigenvectors associated with distinct eigenvalues is a key element in the proof of the following analysis of the spectrum of a linear operator on a nontrivial finite-dimensional Banach space.

3.3.32 Theorem. *Suppose that X is a nontrivial Banach space having finite dimension n and that $T \in B(X)$.*

(a) *The spectrum $\sigma(T)$ of T is finite with cardinality no more than n.*
(b) *$\sigma(T) = \sigma_p(T)$.*
(c) *If α is an eigenvalue of T, then the eigenspaces of T and T^* associated with α have the same dimension.*

PROOF. Part (b) is an immediate consequence of the fact that a member of $B(X)$ is one-to-one if and only if its range is all of X. Part (a) then follows from Theorem 3.3.31. For (c), first note that if $S \in B(X)$, then by Theorem 1.10.17, Lemma 3.1.16, and basic properties of finite-dimensional vector spaces and linear operators on them,

$$\begin{aligned}\dim(\ker(S)) &= \dim\left(\frac{X}{S(X)}\right)\\ &= \dim\left(\left(\frac{X}{S(X)}\right)^*\right)\\ &= \dim(S(X)^\perp)\\ &= \dim(\ker(S^*)).\end{aligned}$$

Letting $S = \alpha I - T$, where α is an eigenvalue of T, completes the proof. ■

In contrast, it is quite possible for the point spectrum of a bounded linear operator from an infinite-dimensional Banach space into itself to be uncountable; see Exercise 3.38. The next section will take up the study of a particular class $K(X, Y)$ of bounded linear operators from a Banach space X into a Banach space Y such that if X is nontrivial, then the spectra of members of $K(X, X)$ have properties very similar to those described in Theorem 3.3.32, even though the Banach space in question might not be finite-dimensional.

Exercises

3.25 The purpose of this exercise is to show that every normed algebra can be completed to a Banach algebra. To this end, suppose that X is a normed algebra. Show that there is a Banach algebra Y and a map T from X onto a dense subspace of Y such that T is a normed algebra isometric isomorphism, that is, a normed space isometric isomorphism for which $T(x_1x_2) = T(x_1)T(x_2)$ whenever $x_1, x_2 \in X$. Show that if X has identity e, then Y has identity $T(e)$.

3.26 Suppose that X is an algebra. Let $X[e]$ be the vector space sum $X \times \mathbb{F}$; see the introduction to Section 1.8 for the definition of a vector space sum. Define a multiplication of elements of $X[e]$ by the formula $(x, \alpha)(y, \beta) = (xy + \alpha y + \beta x, \alpha\beta)$.

(a) Let e be the element $(0, 1)$ of $X[e]$. Prove that with its multiplication and its vector space operations as a vector space sum, the space $X[e]$ is an algebra with identity e.

(b) Suppose that $\|\cdot\|_X$ is a norm for X. Define a norm for $X[e]$ by the formula $\|(x, \alpha)\|_{X[e]} = \|x\|_X + |\alpha|$. Prove that this really does define a norm for $X[e]$. Prove that if X with $\|\cdot\|_X$ is a normed or Banach algebra, then $X[e]$ with $\|\cdot\|_{X[e]}$ is, respectively, a normed or Banach algebra with identity.

(c) With all notation as in (b), suppose that X with the norm $\|\cdot\|_X$ is a normed algebra and therefore that $X[e]$ with the norm $\|\cdot\|_{X[e]}$ is a normed algebra with identity. Prove that the map $T\colon X \to X[e]$ given by the formula $T(x) = (x, 0)$ is a normed algebra isometric isomorphism from X into $X[e]$; see Exercise 3.25 for the definition.

Thus, every Banach algebra without an identity can be embedded into a Banach algebra with identity as a subalgebra having codimension 1.

3.27 Suppose that $t \geq 1$. Give an example of a Banach algebra with identity e such that $\|e\| = t$.

3.28 Suppose that X is a normed algebra with identity e. For each x in X, let $T_x(y) = xy$ whenever $y \in X$. Show that the map $x \mapsto T_x$ is a normed algebra isomorphism (that is, a normed space isomorphism such that $T_{x_1}T_{x_2} = T_{x_1x_2}$ whenever $x_1, x_2 \in X$) from X onto a subalgebra of $B(X)$. Conclude that X is isomorphic as a normed algebra to a normed algebra with identity e' such that $\|e'\| = 1$, whether or not $\|e\| = 1$.

3.29 Suppose that (α_n) is a member of the Banach algebra ℓ_∞. Identify the spectrum of (α_n).

3.30 Let X be the real Banach algebra with identity formed from the complex Banach algebra ℓ_∞ by restricting multiplication of vectors by scalars to $\mathbb{R} \times \ell_\infty$, and let x be the member $(i, 0, 0, \dots)$ of X.

(a) Show that $\sigma(x) = \varnothing$.

(b) Find $\lim_n \|x^n\|^{1/n}$. Comment on what this says about extending the spectral radius formula to real Banach algebras.

3.31 Prove that if X is a complex normed algebra not assumed to be complete or to have an identity, and $x \in X$, then $\lim_n \|x^n\|^{1/n}$ exists. Exercises 3.25 and 3.26 could be helpful.

3.32 Suppose that X is a Banach algebra with identity e and that (x_n) is a sequence of invertible elements of X that converges to some noninvertible x. Prove that $\|x_n^{-1}\| \to \infty$ as $n \to \infty$. (Otherwise, there would be an n such that $\|x_n^{-1}x - e\| < 1$.)

3.33 Prove that if X is a normed algebra with identity, then the map $x \mapsto x^{-1}$ on the set of invertible elements of X is continuous, whether or not X is complete. Exercise 3.25 might help.

3.34 Let e_1 be the element $(1, 0, 0, 0, \dots)$ of the complex Banach algebra ℓ_∞ and let M be the closed subalgebra of ℓ_∞ consisting of all the scalar multiples of e_1, a Banach algebra with identity e_1. Find the spectra of e_1 with respect to the two algebras M and X. Notice that these spectra are different though, of course, the spectral radii are the same.

3.35 A *complex homomorphism* or *multiplicative linear functional* on a complex algebra X is a nonzero linear functional τ on X such that $\tau(xy) = \tau(x)\tau(y)$ whenever $x, y \in X$. Suppose that τ is a complex homomorphism on a complex algebra X with identity e. Prove the following.

(a) $\tau(e) = 1$.

(b) If x is an invertible element of X, then $\tau(x) \neq 0$ and $\tau(x^{-1}) = (\tau(x))^{-1}$.

(c) Suppose that X is a Banach algebra. Then $|\tau(x)| \le \|x\|$ whenever $x \in X$, so $\tau \in X^*$.

3.36 A *division algebra* is an algebra with identity such that the only element of the algebra that is not invertible is 0. Suppose that X is a complex Banach algebra with identity e and that X is also a division algebra. Prove that $X = \{\alpha e : \alpha \in \mathbb{C}\}$. Conclude that X is isomorphic as a normed algebra to $\mathbb{C}$. (See Exercise 3.28 for the definition of a normed algebra isomorphism.) This result is called the *Gelfand-Mazur theorem*.

3.37 An *involution* for an algebra X is a map $x \mapsto x^*$ from X into X such that if $x, y \in X$ and $\alpha \in \mathbb{F}$, then

(1) $(x + y)^* = x^* + y^*$;

(2) $(\alpha x)^* = \overline{\alpha} x^*$;

(3) $(xy)^* = y^* x^*$;

(4) $(x^*)^* = x$.

(It is not being assumed in (2) that $\mathbb{F} = \mathbb{C}$. If $\mathbb{F} = \mathbb{R}$, then $\overline{\alpha}$ is just α.) A Banach algebra X with an involution is a *B^*-algebra* or a *C^*-algebra* if $\|x^*x\| = \|x\|^2$ whenever $x \in X$. Suppose that x is an element of a B^*-algebra X.

(a) Prove that $\|x^*\| = \|x\|$.

Now suppose for the rest of this exercise that the B^*-algebra X has an identity.

(b) Prove that x is invertible if and only if x^* is invertible, in which case $(x^*)^{-1} = (x^{-1})^*$.

(c) Suppose that $\mathbb{F} = \mathbb{C}$ and that x is *self-adjoint,* that is, that $x^* = x$. Prove that $r_\sigma(x) = \|x\|$.

3.38 Let X be complex ℓ_2, let Y be complex ℓ_∞, and let $\mathbb{D}$ be the closed unit disc in the complex plane.

(a) Let T_r be the right-shift operator from X into X given by the formula $T_r(\alpha_1, \alpha_2, \ldots) = (0, \alpha_1, \alpha_2, \ldots)$. Show that $\sigma(T_r) = \mathbb{D}$ but $\sigma_p(T_r) = \varnothing$.

(b) Let T_l be the left-shift operator from X into X given by the formula $T_l(\alpha_1, \alpha_2, \ldots) = (\alpha_2, \alpha_3, \ldots)$. Show that $\sigma(T_l) = \mathbb{D}$ and $\sigma_p(T_l) = \mathbb{D}^\circ$.

(c) Let S_l be the left-shift operator from Y into Y given by the formula for T_l from (b). Show that $\sigma(S_l) = \sigma_p(S_l) = \mathbb{D}$.

3.4 Compact Operators

The following definition could be generalized to linear operators between normed spaces, but enough of the results that follow depend on the completeness of the normed spaces in question to justify concentrating on the Banach space case.

3.4.1 Definition. (D. Hilbert, 1906 [103]; F. Riesz, 1918 [195]). Suppose that X and Y are Banach spaces. A linear operator T from X into Y is *compact* if $T(B)$ is a relatively compact subset of Y whenever B is a bounded subset of X. The collection of all compact linear operators from X into Y is denoted by $K(X, Y)$, or by just $K(X)$ if $X = Y$.

The definition of a compact operator as it appears above is due to Riesz. More will be said about Hilbert's earlier definition near the end of this section.

The notion of a compact operator is in a sense a generalization of the notion of a bounded finite-rank linear operator. To see this, suppose that T is a finite-rank linear operator from a Banach space X into a Banach space Y. Since a subset of $T(X)$ is bounded if and only if it is relatively compact, the operator T is bounded if and only if it is compact. Even if the assumption that $T(X)$ is finite-dimensional is dropped, it is still true that relatively compact subsets of $T(X)$ are bounded, which implies that T is bounded if it is compact. These observations are summarized in the following two results.

3.4.2 Proposition. *Every compact linear operator from a Banach space into a Banach space is bounded.*

3.4.3 Proposition. *A finite-rank linear operator from a Banach space into a Banach space is compact if and only if it is bounded. In particular, a linear functional on a Banach space is compact if and only if it is bounded.*

As will be seen in Example 3.4.5, not every compact operator has finite rank.

Recall that a subset S of a metric space X is *totally bounded* if, for every positive ϵ, there is a finite subset F_ϵ of S (or, equivalently, of X) such that every point of S is within distance ϵ of a member of F_ϵ; that is, such that S is covered by the finite collection of open balls in X having radius ϵ and centered at the points of F_ϵ. It is a basic fact from the theory of metric spaces that a subset of a complete metric space is relatively compact if and only if it is totally bounded; see, for example, [67, p. 22] or [129, p. 101]. This, together with other obvious arguments, proves the following collection of characterizations of compact operators.

3.4.4 Proposition. *Suppose that T is a linear operator from a Banach space X into a Banach space Y. Then the following are equivalent.*

(a) *The operator T is compact.*

(b) *The set $T(B_X)$ is a relatively compact subset of Y.*

(c) *The set $T(B)$ is a totally bounded subset of Y whenever B is a bounded subset of X.*

(d) *Every bounded sequence (x_n) in X has a subsequence (x_{n_j}) such that the sequence (Tx_{n_j}) converges.*

Here is an application of this proposition.

3.4.5 Example. Define the linear operator T from ℓ_2 into itself by the formula $T(\alpha_n) = (n^{-1}\alpha_n)$. Let $\big((\beta_{j,n})\big)_{j=1}^{\infty}$ be a bounded sequence of members of ℓ_2. It will be shown that $\big(T(\beta_{j,n})\big)_{j=1}^{\infty}$ has a convergent subsequence. Since ℓ_2 is reflexive, each of its bounded sequences has a weakly convergent subsequence, so it may be assumed that there is a member (β_n) of ℓ_2 such that $w\text{-}\lim_j(\beta_{j,n}) = (\beta_n)$. It follows that $\lim_j \beta_{j,n} = \beta_n$ for each n. Also, for each positive ϵ there is a positive integer n_ϵ such that $\sum_{n=n_\epsilon}^{\infty} |n^{-1}(\beta_{j,n} - \beta_n)|^2 < \epsilon$ for every j. These two facts together imply that $\lim_j T(\beta_{j,n}) = T(\beta_n)$. It follows from the equivalence of (a) and (d) in Proposition 3.4.4 that T is compact. Notice that $T(X)$ contains the sequence (e_n) of standard unit vectors of ℓ_2, so T is an example of a compact operator that does not have finite rank. Notice also that the range of T is not closed since it contains, for each positive integer n, the element $(1^{-1}, 2^{-1}, \ldots, n^{-1}, 0, 0, \ldots)$ of ℓ_2, but does not contain the limit $(1^{-1}, 2^{-1}, \ldots)$ obtained by letting n tend to ∞.

The failure of the range of T to be closed in the preceding example is a general phenomenon for compact operators not having finite rank.

3.4.6 Proposition. *The range of a compact linear operator from a Banach space into a Banach space is closed if and only if the operator has finite rank.*

PROOF. The reverse implication follows immediately from the fact that finite-dimensional subspaces of normed spaces are always closed. For the forward implication, suppose that T is a compact linear operator from a Banach space X into a Banach space Y having closed range. Then $T(X)$ is itself a Banach space. By the open mapping theorem, the operator T maps open subsets of X onto open subsets of $T(X)$. It follows that T maps the open unit ball of X onto a nonempty relatively compact open subset of $T(X)$, which implies that $T(X)$ has the Heine-Borel property and therefore is finite-dimensional. ■

Of course, the range of a bounded finite-rank linear operator from a Banach space into a Banach space is separable. It turns out that this is a property shared by the ranges of all compact operators.

3.4.7 Proposition. *Every compact linear operator from a Banach space into a Banach space has a separable range.*

PROOF. Suppose that X and Y are Banach spaces and that $T \in K(X,Y)$. Then $[T(B_X)]$ is separable by Theorem 1.12.15 since $T(B_X)$ is relatively compact. It follows that the subspace $T(X)$ of $[T(B_X)]$ is separable. ■

If X and Y are Banach spaces, then $K(X,Y)$ is closed in the usual algebraic and topological senses.

3.4.8 Proposition. *Suppose that X and Y are Banach spaces, that R and S are compact linear operators from X into Y, that $\alpha \in \mathbb{F}$, and that (T_n) is a sequence of compact linear operators from X into Y that converges to some T in $B(X,Y)$. Then $R+S$, αR, and T are all compact.*

PROOF. If (x_n) is a bounded sequence in X, then (x_n) can be thinned to a subsequence (x_{n_j}) such that both (Rx_{n_j}) and (Sx_{n_j}) converge, which implies that $((R+S)(x_{n_j}))$ and (αRx_{n_j}) converge. It follows that $R+S$ and αR are compact.

Now suppose that B is a bounded subset of X and that $\epsilon > 0$. Let n be such that $\|T_n x - Tx\| < \epsilon/3$ whenever $x \in B$. Since $T_n(B)$ is totally bounded, there is a finite subset F of B such that every member of $T_n(B)$ lies within distance $\epsilon/3$ of a member of $T_n(F)$. It follows from straightforward applications of the triangle inequality that every member of $T(B)$ lies within distance ϵ of a member of $T(F)$, which implies that $T(B)$ is totally bounded and therefore that T is compact. ■

With X and Y as in the preceding proposition, the space $K(X,Y)$ is nonempty since it contains the zero operator, so the following corollary is essentially just a restatement of the proposition.

3.4.9 Corollary. *If X and Y are Banach spaces, then $K(X,Y)$ is a closed subspace of $B(X,Y)$.*

It also turns out that the product of two compact operators must be compact. In fact, the compactness of the product follows even if only one of the operators is compact and the other is bounded.

3.4.10 Proposition. *Suppose that X, Y, and Z are Banach spaces, that $S \in B(X,Y)$, and that $T \in B(Y,Z)$. If either S or T is compact, then TS is compact.*

PROOF. By definition, bounded linear operators between normed spaces map bounded sets to bounded sets, and also map relatively compact sets to relatively compact sets because of their continuity. The proposition follows easily from this. ■

When A is a subset of an algebra X and $x \in X$, the sets $\{\, xy : y \in A \,\}$ and $\{\, yx : y \in A \,\}$ are denoted, as one might expect, by xA and Ax respectively.

3.4.11 Definition. An *ideal* in an algebra X is a vector subspace M of X such that $xM, Mx \subseteq M$ whenever $x \in X$.

It is clear that every ideal in an algebra is a subalgebra of the algebra. With this term in hand, and in light of Corollary 3.4.9, the following result is an immediate corollary of Proposition 3.4.10.

3.4.12 Corollary. *If X is a Banach space, then $K(X)$ is a closed ideal in $B(X)$.*

For a bounded linear operator between Banach spaces, the property of being compact can be characterized in several different ways in terms of the behavior of the adjoint of the operator. One such characterization will be derived here from the Arzelà-Ascoli theorem, and a second will then be obtained from the first.

3.4.13 Definition. Suppose that K is a compact Hausdorff space and that S is a subset of $C(K)$ with the property that for every x in K and every positive ϵ there is a neighborhood $U_{x,\epsilon}$ of x such that $|f(y) - f(x)| < \epsilon$ whenever $f \in S$ and $y \in U_{x,\epsilon}$. Then S is *equicontinuous.*

In the following theorem, Arzelà proved the forward implication while Ascoli is responsible for the converse.

3.4.14 The Arzelà-Ascoli Theorem. (C. Arzelà, 1889 [6]; G. Ascoli, 1882–1883 [7]). *Suppose that K is a compact Hausdorff space and that S is a subset of $C(K)$. Then S is relatively compact if and only if S is bounded and equicontinuous.*

PROOF. It may be assumed that K and S are nonempty. Suppose first that S is relatively compact. Fix a positive ϵ and an x_0 in K. Since S is totally bounded, there is a finite subset F of S such that every member of S lies within distance $\epsilon/3$ of a member of F. Let $U_{x_0,\epsilon}$ be a neighborhood of x_0 such that $|f(y) - f(x_0)| < \epsilon/3$ whenever $f \in F$ and $y \in U_{x_0,\epsilon}$. It follows that if $f_0 \in S$ and $y \in U_{x_0,\epsilon}$, then there is an f in F such that

$$|f_0(y) - f_0(x_0)| \leq |f_0(y) - f(y)| + |f(y) - f(x_0)| + |f(x_0) - f_0(x_0)| < \epsilon.$$

The set S is therefore equicontinuous, and is bounded since it is relatively compact.

Suppose conversely that S is bounded and equicontinuous. Again, fix a positive ϵ. For each x in K, let $U_{x,\epsilon}$ be a neighborhood of x such that $|f(y) - f(x)| < \epsilon/3$ whenever $f \in S$ and $y \in U_{x,\epsilon}$. Since K is compact, there is a finite collection $x_1, \ldots, x_m$ of members of K such that $K \subseteq \bigcup_{j=1}^{m} U_{x_j,\epsilon}$. It follows from the boundedness of S that $\{\, (f(x_1), \ldots, f(x_m)) : f \in S \,\}$ is a subset of Euclidean n-space that is bounded and therefore totally bounded, so there is a finite collection $f_1, \ldots, f_n$ of members of S such that for each f in S there is a j for which $(f(x_1), \ldots, f(x_m))$ is within distance $\epsilon/3$ of $(f_j(x_1), \ldots, f_j(x_m))$. Fix an f in S and let f_j be as in the preceding sentence. If $x \in K$, then there is a k such that $x \in U_{x_k,\epsilon}$, and so

$$|f(x) - f_j(x)| \leq |f(x) - f(x_k)| + |f(x_k) - f_j(x_k)| + |f_j(x_k) - f_j(x)| < \epsilon.$$

It follows that $\|f - f_j\|_\infty < \epsilon$. The set S is therefore totally bounded and so is relatively compact. ■

The following characterization of operator compactness is important enough to have inherited the name of its discoverer.

3.4.15 Schauder's Theorem. (J. Schauder, 1930 [209]). *A bounded linear operator from a Banach space into a Banach space is compact if and only if its adjoint is compact.*

PROOF. Let T be a bounded linear operator from a Banach space X into a Banach space Y. Suppose first that T is compact. Let $K = \overline{T(B_X)}$ and let B be a bounded subset of Y^*. Since

$$|z^*y_1 - z^*y_2| \leq \|y_1 - y_2\| \sup\{\, \|y^*\| : y^* \in B \,\}$$

whenever $z^* \in B$ and $y_1, y_2 \in K$, the set B, viewed as a subset of $C(K)$, is bounded and equicontinuous and therefore relatively compact by the

Arzelà-Ascoli theorem. Suppose that (y_n^*) is a sequence in B. By what has been proved, some subsequence $(y_{n_j}^*)$ of (y_n^*) is uniformly Cauchy on K, which implies that $(y_{n_j}^* T)$ is uniformly Cauchy on B_X and therefore, by the completeness of X^*, convergent as a sequence in X^*. Since $\big(T^*(y_{n_j}^*)\big) = (y_{n_j}^* T)$, it follows that T^* is compact.

Conversely, suppose that T^* is compact. Then T^{**} is compact by what has already been proved. Let Q_X and Q_Y be the natural maps from X and Y respectively into their second duals. Then $T = Q_Y^{-1} T^{**} Q_X$ by Proposition 3.1.13, from which it follows that T is compact. ■

As was promised earlier, there is another important characterization of operator compactness in terms of the behavior of the adjoint, and it is based on a continuity condition involving the bounded weak* topology.

3.4.16 Theorem. *A bounded linear operator from a Banach space X into a Banach space Y is compact if and only if its adjoint is continuous from the bounded weak* topology of Y^* to the norm topology of X^*.*

PROOF. Let T be a bounded linear operator from X into Y. Suppose first that T is compact. Then T^* is also compact. Let F be a closed subset of X^*. To show that T^* has the claimed continuity property, it is enough to show that $(T^*)^{-1}(F)$ is b-weakly* closed. Let (y_α^*) be a bounded net in $(T^*)^{-1}(F)$ that is weakly* convergent to some y^* in Y^*. By Corollary 2.7.5, the set $(T^*)^{-1}(F)$ will be shown to be b-weakly* closed once it is shown that $y^* \in (T^*)^{-1}(F)$. Now the weak*-to-weak* continuity of T^* assures that $T^* y_\alpha^* \xrightarrow{w^*} T^* y^*$, while the compactness of T^* produces a subnet (y_β^*) of (y_α^*) such that $(T^* y_\beta^*)$ converges in the norm topology to some x^* in X^* that must lie in the closed set F. It follows that $T^* y^* = x^*$, so $y^* \in (T^*)^{-1}(F)$ and T^* has the required continuity property.

Suppose conversely that T^* is continuous from the bounded weak* topology of Y^* to the norm topology of X^*. Let B be a bounded subset of Y^* and let (z_γ^*) be a net in B. To show that T^* is compact and therefore that T is compact, it is enough to show that $(T^* z_\gamma^*)$ has a subnet convergent in the norm topology. By the Banach-Alaoglu theorem, the bounded net (z_γ^*) has a subnet (z_δ^*) that is convergent in the weak* topology and therefore in the bounded weak* topology, which implies that $(T^* z_\delta^*)$ is convergent in the norm topology. ■

3.4.17 Corollary. *If the adjoint of a bounded linear operator T from a Banach space into a Banach space is weak*-to-norm continuous, then T is compact.*

It is not true that every compact operator has a weak*-to-norm continuous adjoint. See Exercise 3.44.

The next portion of this section is devoted to obtaining Theorem 3.4.23, Frédéric Riesz's analysis of the spectrum of a compact operator from an infinite-dimensional Banach space into itself. The path that will be taken to this theorem is related to Riesz's original method. See [67, p. 609] for the history of Riesz's result and its relationship to earlier work of Fredholm, and for references to various methods of obtaining it.

The first lemma en route to Theorem 3.4.23 is so useful throughout functional analysis that it has become a named result.

3.4.18 Riesz's Lemma. (F. Riesz, 1918 [195]). *Suppose that M is a proper closed subspace of a normed space X and that $0 < \theta < 1$. Then there is an x in S_X such that $d(x, M) \geq \theta$.*

PROOF. By Corollary 1.9.7, there is an x^* in S_{X^*} such that $M \subseteq \ker(x^*)$. Let x be an element of S_X such that $|x^*x| \geq \theta$. If $y \in M$, then $\|x - y\| \geq |x^*x - x^*y| = |x^*x| \geq \theta$, as required. ■

Suppose that X is a Banach space. It often happens that if a statement is true for $I - T$ whenever $T \in K(X)$, then it must also be true for $\alpha I - T$ whenever $T \in K(X)$ and α is a nonzero scalar since $\alpha^{-1}T$ is compact and $\alpha I - T = \alpha(I - \alpha^{-1}T)$. This can sometimes be used to simplify the notation in certain arguments just a bit, as is done in the proofs of the next three lemmas.

3.4.19 Lemma. *Suppose that X is a Banach space, that $T \in K(X)$, and that α is a nonzero scalar. If $(\alpha I - T)(X) = X$, then $\alpha I - T$ is one-to-one.*

PROOF. It may be assumed that $\alpha = 1$. Notice that T commutes with $(I - T)^n$ for each positive integer n; think of the expansion of $(I - T)^n$ into polynomial form. It follows that

$$T\Big(\ker\big((I - T)^n\big)\Big) \subseteq \ker\big((I - T)^n\big)$$

whenever $n \in \mathbb{N}$.

Suppose that $(I-T)(X) = X$ but $I-T$ is not one-to-one. If n is a positive integer, then $(I-T)^n(X) = X$ and therefore $(I-T)^{n+1}$ maps some member of X to 0 that $(I - T)^n$ does not, so $\ker\big((I - T)^n\big) \subsetneqq \ker\big((I - T)^{n+1}\big)$. By Riesz's lemma, there is a sequence (x_n) in S_X such that if $n \in \mathbb{N}$ and $y \in \ker\big((I - T)^n\big)$, then $x_n \in \ker\big((I - T)^{n+1}\big)$ and $\|x_n - y\| \geq 1/2$. If $n, m \in \mathbb{N}$ and $n > m$, then $(I - T)(x_n) + Tx_m \in \ker\big((I - T)^n\big)$, so

$$\|Tx_n - Tx_m\| = \Big\|x_n - \big((I - T)(x_n) + Tx_m\big)\Big\| \geq \frac{1}{2}.$$

It follows that (Tx_n) has no convergent subsequence, which contradicts the compactness of T. ■

Under the hypotheses of the preceding lemma, it actually turns out that $(\alpha I - T)(X) = X$ if and only if $\alpha I - T$ is one-to-one. This is an immediate consequence of the next lemma.

3.4.20 Lemma. *Suppose that X is a Banach space, that $T \in K(X)$, and that α is a nonzero scalar. Then $\alpha I - T$ has finite-dimensional kernel and closed range, and the dimension of its kernel equals the codimension of its range.*

PROOF. It may be assumed that $\alpha = 1$. If $\ker(I - T)$ were infinite-dimensional, then, by Lemma 1.4.22, the unit sphere of $\ker(I - T)$ would contain a sequence (x_n) such that $\|x_n - x_m\| \geq 1$ whenever $n \neq m$, which would imply that $\|Tx_n - Tx_m\| = \|x_n - x_m\| \geq 1$ whenever $n \neq m$, and so (Tx_n) would have no convergent subsequence. This contradicts the compactness of T, so $\ker(I - T)$ must be finite-dimensional.

It follows from Theorem 3.2.18 (b) that $\ker(I - T)$ is a complemented subspace of X. Let M be a closed subspace of X such that X is the internal direct sum of $\ker(I - T)$ and M. Then $(I - T)(M) = (I - T)(X)$ and $I - T$ is one-to-one on M. To show that $(I - T)(X)$ is closed, it is enough to show that the restriction of $I - T$ to M is an isomorphism. Suppose it were not. Then there would be a sequence (y_n) in S_M such that $(I - T)(y_n) \to 0$. Due to the compactness of T, it may be assumed that (Ty_n) converges to some z, which implies that $y_n = (I - T)(y_n) + Ty_n \to z$ and therefore that $z \in S_M$. However,

$$(I - T)(z) = \lim_n (I - T)(y_n) = 0,$$

which implies that $z = 0$ since $I - T$ is one-to-one on M. This contradicts the fact that $z \in S_M$ and therefore establishes that $(I - T)(X)$ is closed.

It remains to be shown that $\dim(\ker(I - T)) = \operatorname{codim}((I - T)(X))$. The first step toward this is to observe that, by Theorem 1.10.17 and Lemma 3.1.16,

$$\left(\frac{X}{(I - T)(X)}\right)^* \cong ((I - T)(X))^\perp = \ker(I^* - T^*).$$

Since $\ker(I^* - T^*)$ is finite-dimensional by the compactness of T^* and since a finite-dimensional normed space has the same dimension as its dual space,

$$\operatorname{codim}((I - T)(X)) = \dim(\ker(I^* - T^*)). \tag{3.4}$$

The compactness of T^* also implies that $(I^* - T^*)(X^*)$ is closed and thus, by Theorem 3.1.21, weakly* closed. It follows from Lemma 3.1.16 and Proposition 2.6.6 (c) that

$$(\ker(I - T))^\perp = \left({}^\perp((I^* - T^*)(X^*))\right)^\perp = (I^* - T^*)(X^*)$$

and therefore from Theorem 1.10.16 that

$$\frac{X^*}{(I^* - T^*)(X^*)} = \frac{X^*}{(\ker(I - T))^\perp} \cong (\ker(I - T))^*.$$

A further application of the fact that every finite-dimensional normed space has the same dimension as its dual then shows that

$$\operatorname{codim}\big((I^* - T^*)(X^*)\big) = \dim\big(\ker(I - T)\big). \tag{3.5}$$

Next, suppose that $\dim\big(\ker(I - T)\big) > \operatorname{codim}\big((I - T)(X)\big)$. It follows from Theorem 3.2.18 (a) that there is a closed subspace N of X such that X is the internal direct sum of $(I - T)(X)$ and N. Applying the first isomorphism theorem for Banach spaces to the projection with range N and kernel $(I - T)(X)$ shows that $\operatorname{codim}\big((I - T)(X)\big) = \dim(N)$, so there is a bounded linear operator S that is not one-to-one that maps the finite-dimensional Banach space $\ker(I - T)$ onto the smaller-dimensional Banach space N. Consider S to be a member of $B(\ker(I - T), X)$, and notice that S is compact. As in the first part of this proof, let M be a closed subspace of X such that X is the internal direct sum of $\ker(I - T)$ and M, and let P be the projection in X with range $\ker(I - T)$ and kernel M. Let $R = T + SP$. Then R is compact since T and S are so, and

$$\begin{aligned}(I - R)(X) &= \big((I - T) - SP\big)\big(\ker(I - T)\big) + \big((I - T) - SP\big)(M)\\ &= N + (I - T)(M)\\ &= N + (I - T)(X)\\ &= X.\end{aligned}$$

It follows from Lemma 3.4.19 that $I - R$ is one-to-one. However, there is some nonzero w in the subset $\ker(S)$ of $\ker(I - T)$, which implies that $(I - R)(w) = (I - T)(w) - SPw = 0$, a contradiction. Therefore, it must be that

$$\dim\big(\ker(I - T)\big) \le \operatorname{codim}\big((I - T)(X)\big), \tag{3.6}$$

and furthermore, since T^* is compact, that

$$\dim\big(\ker(I^* - T^*)\big) \le \operatorname{codim}\big((I^* - T^*)(X^*)\big). \tag{3.7}$$

It follows from (3.4), (3.5), (3.6), and (3.7) that

$$\dim\big(\ker(I - T)\big) = \operatorname{codim}\big((I - T)(X)\big),$$

as required. ∎

The next result also follows from (3.4), (3.5), (3.6), and (3.7).

3.4.21 Lemma. *Suppose that X is a Banach space, that $T \in K(X)$, and that α is a nonzero scalar. Then $\dim\big(\ker(\alpha I - T)\big) = \dim\big(\ker(\alpha I^* - T^*)\big)$.*

3.4.22 Lemma. *Suppose that X is a nontrivial Banach space, that $T \in K(X)$, and that (α_n) is a sequence of distinct eigenvalues of T. Then $\alpha_n \to 0$.*

PROOF. It may be assumed that each α_n is nonzero. For each positive integer n, let x_n be an eigenvector associated with α_n and let $M_n = \langle\{x_1, \ldots, x_n\}\rangle$, a closed subspace of X that has dimension n by Theorem 3.3.31; then $M_n \subsetneqq M_{n+1}$ and $(\alpha_{n+1}I - T)(M_{n+1}) \subseteq M_n$, so by Riesz's lemma there is a y_{n+1} in $S_{M_{n+1}}$ such that $d(y_{n+1}, M_n) \geq 1/2$ and $(\alpha_{n+1}I - T)(y_{n+1}) \in M_n$. It follows that if $j, k \in \mathbb{N}$ and $j > k > 1$, then

$$\begin{aligned}T(\alpha_j^{-1}y_j) - T(\alpha_k^{-1}y_k)\\ &= \alpha_k^{-1}(\alpha_k I - T)(y_k) - \alpha_j^{-1}(\alpha_j I - T)(y_j) - y_k + y_j\\ &= y_j - \big(y_k + \alpha_j^{-1}(\alpha_j I - T)(y_j) - \alpha_k^{-1}(\alpha_k I - T)(y_k)\big).\end{aligned}$$

The expression subtracted from y_j lies in M_{j-1} and therefore has distance at least 1/2 from y_j, so $\|T(\alpha_j^{-1}y_j) - T(\alpha_k^{-1}y_k)\| \geq 1/2$. Therefore $\big(T(\alpha_n^{-1}y_n)\big)$ has no convergent subsequence. Since T is compact, it must be that $(\alpha_n^{-1}y_n)$ has no bounded subsequence, which implies that $\alpha_n \to 0$. ∎

3.4.23 Theorem. (F. Riesz, 1918 [195]). *Suppose that X is an infinite-dimensional Banach space and that $T \in K(X)$.*

(a) *The spectrum $\sigma(T)$ of T is a countable compact set whose only possible limit point is 0.*

(b) $\sigma(T) = \{0\} \cup \sigma_p(T)$.

(c) *If α is a nonzero eigenvalue of T, then the eigenspaces of T and T^* associated with α have the same finite dimension.*

PROOF. Since X is infinite-dimensional, the compact operator T cannot be an isomorphism from X onto itself, so $0 \in \sigma(T)$. If $\alpha \in \sigma(T) \setminus \{0\}$, then either $\ker(\alpha I - T) \neq \{0\}$ or $(\alpha I - T)(X) \neq X$; it follows from Lemma 3.4.20 that both are actually true, so $\alpha \in \sigma_p(T)$. This proves (b).

It follows from Theorem 3.3.18 that $\sigma(T)$ is compact. Now suppose that $r > 0$, that $A_r = \{\,\alpha : \alpha \in \sigma(T),\ |\alpha| \geq r\,\}$, and that A_r is infinite. Then some sequence of distinct members of A_r converges to some member of A_r. By (b), each member of this sequence is an eigenvalue of T, so this sequence converges to 0 by Lemma 3.4.22. This contradiction assures that A_r is finite for each positive r, from which (a) follows.

Now suppose that α is a nonzero eigenvalue of T. Then $\dim\big(\ker(\alpha I - T)\big)$ is finite by Lemma 3.4.20 and equal to $\dim\big(\ker(\alpha I^* - T^*)\big)$ by Lemma 3.4.21, which proves (c). ∎

The preceding theorem does not say that the spectrum of a compact operator from an infinite-dimensional Banach space into itself is the *disjoint* union of $\{0\}$ and $\sigma_p(T)$. For example, the compact operator $(\alpha_n) \mapsto (\alpha_1, 0, 0, \ldots)$ from ℓ_2 into itself has 0 as an eigenvalue (and the associated eigenspace is infinite-dimensional). On the other hand, notice that the member T of $K(\ell_2)$ from Example 3.4.5 is one-to-one and therefore does not have 0 as an eigenvalue.

Several of the lemmas leading up to the preceding theorem have an interesting application to the solution of linear equations.

3.4.24 The Fredholm Alternative. *Suppose that X is a Banach space, that $T \in K(X)$, and that α is a nonzero scalar. Then the following are equivalent.*

(a) *For every y in X, the equation $(\alpha I - T)(x) = y$ is satisfied by some x in X.*

(b) *For every y^* in X^*, the equation $(\alpha I^* - T^*)(x^*) = y^*$ is satisfied by some x^* in X^*.*

(c) *The homogeneous equation $(\alpha I - T)(x) = 0$ is satisfied only when $x = 0$.*

(d) *The homogeneous equation $(\alpha I^* - T^*)(x^*) = 0$ is satisfied only when $x^* = 0$.*

If any (and therefore all) of (a), (b), (c), *and* (d) *are true, then the equation in* (a) *has a unique solution x_y whenever $y \in X$, and similarly for the equation in* (b). *Whether or not* (c) *and* (d) *are true, the solution spaces of the homogeneous equations in* (c) *and* (d) *have the same finite dimension.*

PROOF. The fact that the solution spaces of the homogeneous linear equations in (c) and (d) have the same finite dimension is just a restatement of Lemma 3.4.21 and part of Lemma 3.4.20. The equivalence of (c) and (d) follows from this. It is a consequence of Lemma 3.4.20 that $\alpha I - T$ is one-to-one if and only if $(\alpha I - T)(X) = X$, which is another way to say that (a) and (c) are equivalent; this also shows that if (a) holds, then the equation in (a) is satisfied by a unique x_y whenever $y \in X$. The same argument applied to T^* proves the equivalence of (b) and (d) and the uniqueness assertion about solutions of the equation in (b). ■

See Exercise 3.47 for more properties of solutions of homogeneous and nonhomogeneous equations that are sometimes included in the Fredholm alternative.

Most of the rest of this section is devoted to a very brief look at the approximation property. Historically, the motivating example is the following one.

3.4.25 Example. For each positive integer n, define $P_n \colon \ell_2 \to \ell_2$ by the formula $P_n\big((\alpha_j)\big) = (\alpha_1, \dots, \alpha_n, 0, 0, \dots)$. Then each P_n is a bounded projection from ℓ_2 onto $\{\,(\alpha_j) : (\alpha_j) \in \ell_2,\ \alpha_{n+1} = \alpha_{n+2} = \cdots = 0\,\}$. Suppose that S is a nonempty relatively compact subset of ℓ_2. For each positive ϵ, there must be a positive integer n_ϵ such that $\big(\sum_{j=n_\epsilon+1}^{\infty} |\alpha_j|^2\big)^{1/2} < \epsilon$ whenever $(\alpha_j) \in S$, because if this failed for some positive ϵ then it would be possible to construct a sequence of elements of S with no convergent subsequence; the basic idea is to start with some $(\alpha_{1,j})$ in S, select an n_1 such that $(\sum_{j=n_1+1}^{\infty}|\alpha_{1,j}|^2)^{1/2} \leq \epsilon/2$, then select an $(\alpha_{2,j})$ in S such that $(\sum_{j=n_1+1}^{\infty}|\alpha_{2,j}|^2)^{1/2} \geq \epsilon$, and continue in the obvious fashion. It follows that $\lim_n\{\sup\{\, \|(I - P_n)(y)\|_2 : y \in S\,\}\} = 0$.

Now suppose that T is a compact operator from a Banach space X into ℓ_2. Let $T_n = P_nT$ for each n. Then each T_n is a bounded finite-rank linear operator, and

$$\begin{aligned} 0 &= \lim_n\{\sup\{\, \|(I - P_n)(y)\|_2 : y \in T(B_X)\,\}\} \\ &= \lim_n\{\sup\{\, \|(T - T_n)(x)\|_2 : x \in B_X\,\}\} \\ &= \lim_n \|T - T_n\|. \end{aligned}$$

This proves that *every compact linear operator from a Banach space into ℓ_2 is the limit of a sequence of bounded finite-rank linear operators from that Banach space into ℓ_2.* Since bounded finite-rank linear operators from a Banach space into a Banach space are compact, another way of expressing this is to say that ℓ_2 has the following property.

3.4.26 Definition. A Banach space X has the *approximation property* if, for every Banach space Y, the set of finite-rank members of $B(Y, X)$ is dense in $K(Y, X)$.

The argument that proves the following result is essentially the same as that of Example 3.4.25.

3.4.27 Theorem. *The spaces c_0 and ℓ_p such that $1 \leq p < \infty$ have the approximation property.*

Though the definition of the approximation property given above is the historical one, it is probably not the one used most often by modern authors. The definition commonly encountered today is an equivalent one due to Alexandre Grothendieck that has the advantage of being based on intrinsic properties of the Banach space X in question rather than on properties that involve *every* Banach space Y as well as X. The equivalence of these two definitions will be proved in Theorem 3.4.32, both for the importance of the result and because some useful ideas occur in the proofs of the theorem and the lemmas leading up to it.

3.4.28 Lemma. *Suppose that S is a nonempty relatively compact subset of a Banach space X and that $\|\cdot\|_X$ is the norm of X. Let*

$$K_S = \overline{\mathrm{co}}\left(\bigcup\{\,\alpha S : \alpha \in \mathbb{F},\ |\alpha| \leq 1\,\}\right)$$

and let $Y = \langle K_S \rangle$. Then

(a) *the set K_S is a compact subset of X that includes S;*

(b) *the vector space Y has a Banach norm $\|\cdot\|_Y$ such that K_S is the closed unit ball for that norm; and*

(c) *the "identity" map from $(Y, \|\cdot\|_Y)$ into $(X, \|\cdot\|_X)$ is compact.*

PROOF. The set $\bigcup\{\,\alpha S : \alpha \in \mathbb{F},\ |\alpha| \le 1\,\}$ is relatively compact because each of its sequences has a convergent subsequence. Since

$$K_S = \overline{\mathrm{co}}\left(\overline{\bigcup\{\,\alpha S : \alpha \in \mathbb{F},\ |\alpha| \le 1\,\}}\right),$$

the set K_S is compact by Mazur's compactness theorem. This and the obvious fact that $S \subseteq K_S$ prove (a).

Since $\mathrm{co}(\bigcup\{\,\alpha S : \alpha \in \mathbb{F},\ |\alpha| \le 1\,\})$ is balanced, so is K_S by Theorem 2.2.9 (i), which implies that K_S is an absorbing subset of Y since $Y = \langle K_S \rangle$ and K_S is convex. Let $\|\cdot\|_Y$ be the Minkowski functional of K_S on Y, a seminorm by Proposition 1.9.14 (a) (3). If y is a nonzero member of Y, then $\|y\|_Y > 0$, for otherwise there would be a sequence t_n of positive reals converging to 0 such that $t_n^{-1}y \in K_S$ for all n, contradicting the boundedness of K_S in X. Therefore $\|\cdot\|_Y$ is a norm on Y. It is easy to check that $\|y\|_Y \le 1$ if $y \in K_S$ and $\|y\|_Y > 1$ if $y \in Y \setminus K_S$, that is, that K_S is the closed unit ball for $\|\cdot\|_Y$. The "identity" map from $(Y, \|\cdot\|_Y)$ into $(X, \|\cdot\|_X)$ is bounded since it maps $B_{(Y,\|\cdot\|_Y)}$ onto the bounded subset K_S of X, and will necessarily be compact once it is demonstrated that $(Y, \|\cdot\|_Y)$ is a Banach space.

All that remains to be proved is that the metric induced by $\|\cdot\|_Y$ is complete. Suppose it is not. Let (v_n) be a nonconvergent Cauchy sequence in $B_{(Y,\|\cdot\|_Y)}$. Since (v_n) is also Cauchy in X and lies in K_S, there is some v in K_S such that $\|v_n - v\|_X \to 0$. Letting $w_n = v_n - v$ for each n produces a nonconvergent Cauchy sequence in Y such that $\|w_n\|_X \to 0$. It follows that there is a subsequence (w_{n_j}) of (w_n) and a positive δ such that $\|w_{n_j}\|_Y \ge \delta$ for each j. Letting $z_j = \|w_{n_j}\|_Y^{-1} w_{n_j}$ for each j produces a Cauchy sequence in $S_{(Y,\|\cdot\|_Y)}$ that, when viewed as a sequence in X, converges to 0. Let j_0 be a positive integer such that $\|z_j - z_k\|_Y \le 1/2$ whenever $j, k \ge j_0$. Then $2(z_{j_0} - z_j) \in K_S$ whenever $j \ge j_0$. Since $z_j \to 0$ in X and K_S is closed in X, it follows that $2z_{j_0} \in K_S$. This implies that $\|z_{j_0}\|_Y \le 1/2$, which is a contradiction. ■

3.4.29 Lemma. *Suppose that (x_n) is a sequence in a Banach space X that converges to 0 and that the subset H of $\mathbb{F}$ is either $\{1\}$ or a closed ball in $\mathbb{F}$ centered at 0. Then*

$$\overline{\mathrm{co}}(\{\,\alpha x_n : \alpha \in H,\ n \in \mathbb{N}\,\})$$
$$= \left\{\sum_n t_n \alpha_n x_n : t_n \ge 0 \text{ and } \alpha_n \in H \text{ for each } n,\ \sum_n t_n \le 1\right\}$$

and this closed convex hull is compact.

PROOF. The set $\{\,\alpha x_n : \alpha \in H,\ n \in \mathbb{N}\,\} \cup \{0\}$ is compact because each sequence in it has a convergent subsequence whose limit is in the set. Also,

$$\overline{\mathrm{co}}(\{\,\alpha x_n : \alpha \in H,\ n \in \mathbb{N}\,\}) = \overline{\mathrm{co}}(\{\,\alpha x_n : \alpha \in H,\ n \in \mathbb{N}\,\} \cup \{0\})$$

and this last set is compact by Mazur's compactness theorem. Let

$$B = \Big\{ \sum_n t_n \alpha_n x_n : t_n \ge 0 \text{ and } \alpha_n \in H \text{ for each } n,\ \sum_n t_n \le 1 \Big\}.$$

Notice that there is no problem with the definition of B, since X is a Banach space and each of the formal series that is supposed to belong to B is absolutely convergent. It will now be shown that B is closed. Suppose that $(\sum_n t_{j,n}\alpha_{j,n}x_n)_{j=1}^\infty$ is a sequence in B convergent to some member of X. By a straightforward diagonalization argument, it may be assumed that, for each n, the sequence $(t_{j,n})_{j=1}^\infty$ converges to some nonnegative t_n and $(\alpha_{j,n})_{j=1}^\infty$ converges to some α_n in H. It follows that $\sum_n t_n \le 1$ and, by an argument involving the fact that $x_n \to 0$, that $\lim_j \sum_n t_{j,n}\alpha_{j,n}x_n = \sum_n t_n\alpha_n x_n$. Therefore B is closed. It is not difficult to show that B is convex; when H is a closed ball in $\mathbb{F}$ centered at 0, this uses the fact that if $\sum_n t_n\alpha_n x_n, \sum_n t'_n\alpha'_n x_n \in B$ and $0 < t < 1$, then

$$|\, t t_n \alpha_n + (1-t)t'_n\alpha'_n| \le \big(t t_n + (1-t)t'_n\big) \max\{\, |\alpha| : \alpha \in H\,\}$$

for each n. Since

$$\{\,\alpha x_n : \alpha \in H,\ n \in \mathbb{N}\,\} \subseteq B \subseteq \overline{\mathrm{co}}(\{\,\alpha x_n : \alpha \in H,\ n \in \mathbb{N}\,\}),$$

it follows that $B = \overline{\mathrm{co}}(\{\,\alpha x_n : \alpha \in H,\ n \in \mathbb{N}\,\})$. ■

3.4.30 Lemma. *Suppose that A is a relatively compact subset of a normed space X. Then there is a sequence (x_n) in X converging to 0 such that $A \subseteq \overline{\mathrm{co}}(\{\, x_n : n \in \mathbb{N}\,\})$.*

PROOF. It may be assumed that $A \neq \emptyset$. For each x in X and each positive r, let $B(x,r)$ denote the closed ball of radius r centered at x. Since $2A$ is relatively compact and therefore totally bounded, there are members $x_1, \dots, x_{n_1}$ of $2A$ such that $2A \subseteq \bigcup_{j=1}^{n_1} B(x_j, 2^{-1})$. Let

$$A_1 = \bigcup_{j=1}^{n_1} \Big((2A \cap B(x_j, 2^{-1})) - x_j \Big).$$

Then $A_1 \subseteq B(0, 2^{-1})$. Since A_1 is nonempty and relatively compact, there are members $x_{n_1+1}, \dots, x_{n_2}$ of $2A_1$ such that $2A_1 \subseteq \bigcup_{j=n_1+1}^{n_2} B(x_j, 2^{-2})$. Let

$$A_2 = \bigcup_{j=n_1+1}^{n_2} \Big((2A_1 \cap B(x_j, 2^{-2})) - x_j \Big).$$

Then $A_2 \subseteq B(0, 2^{-2})$. Since A_2 is nonempty and relatively compact, there are members $x_{n_2+1}, \ldots, x_{n_3}$ of $2A_2$ such that $2A_2 \subseteq \bigcup_{j=n_2+1}^{n_3} B(x_j, 2^{-3})$. Let

$$A_3 = \bigcup_{j=n_2+1}^{n_3} \Big((2A_2 \cap B(x_j, 2^{-3})) - x_j \Big).$$

The construction of (x_n) is continued in the obvious fashion. Notice that $x_n \to 0$.

Suppose that $x_0 \in A$. Then there is a positive integer j_1 with $1 \le j_1 \le n_1$ such that $2x_0 - x_{j_1} \in A_1$, so there is a positive integer j_2 with $n_1 < j_2 \le n_2$ such that $4x_0 - 2x_{j_1} - x_{j_2} \in A_2$, and so forth. It follows that

$$x_0 - \sum_{n=1}^{m} 2^{-n} x_{j_n} \in 2^{-m} A_m \subseteq B(0, 4^{-m})$$

for each m, and therefore that $x_0 = \sum_n 2^{-n} x_{j_n} \in \overline{\mathrm{co}}(\{ x_n : n \in \mathbb{N} \})$. ∎

Thus, every relatively compact subset of a normed space is small in the sense that it is included in the closed convex hull of a null sequence. If the normed space is not a Banach space, then such a closed convex hull is not guaranteed to be compact; see Exercise 2.94. However, it is when the space is complete, as was shown in Lemma 3.4.29.

3.4.31 Lemma. *Suppose that X is a Banach space.*

(a) *A subset A of X is relatively compact if and only if there is a sequence (x_n) in X converging to 0 such that $A \subseteq \overline{\mathrm{co}}(\{ x_n : n \in \mathbb{N} \})$.*

(b) *If A is a relatively compact subset of X and (x_n) is as in (a), then there is a compact subset S of X such that $\overline{\mathrm{co}}(\{ x_n : n \in \mathbb{N} \}) \subseteq S$ and, with Y being the Banach space constructed from this S as in Lemma 3.4.28, the sequence (x_n) converges to 0 in Y as well as in X.*

PROOF. The forward implication in (a) follows immediately from the preceding lemma, while the converse comes from Lemma 3.4.29 and the fact that every subset of a compact subset of X is relatively compact.

For (b), suppose that A is a relatively compact subset of X and that (x_n) is a sequence in X as in (a). It may be assumed that each x_n is nonzero. Let $\|\cdot\|_X$ be the norm of X. For each n, let $y_n = \|x_n\|_X^{-1/2} x_n$ if $\|x_n\|_X < 1$ and let $y_n = x_n$ otherwise. Let $S = \overline{\mathrm{co}}(\{ y_n : n \in \mathbb{N} \})$, a compact subset of X since $\|y_n\|_X \to 0$. Then $\overline{\mathrm{co}}(\{ x_n : n \in \mathbb{N} \}) \subseteq S$. Let Y and $\|\cdot\|_Y$ be as in Lemma 3.4.28 for this S. If $\|x_n\|_X < 1$, then $\|x_n\|_X^{-1/2} x_n \in B_{(Y, \|\cdot\|_Y)}$, so $\|x_n\|_Y \le \|x_n\|_X^{1/2}$. It follows that $\|x_n\|_Y \to 0$, which proves (b). ∎

3.4.32 Theorem. (A. Grothendieck, 1955 [96]). *Let X be a Banach space. Then the following are equivalent.*

(a) *The space X has the approximation property.*

(b) *For every compact subset K of X and every positive ϵ, there is a finite-rank member $T_{K,\epsilon}$ of $B(X)$ such that $\|T_{K,\epsilon}x - x\| < \epsilon$ whenever $x \in K$. That is, the identity operator on X can be uniformly approximated on compact subsets of X by bounded finite-rank linear operators.*

PROOF. Suppose first that X has the approximation property, that K is a compact subset of X, and that $\epsilon > 0$. By Lemma 3.4.31, there is a sequence (x_n) in X converging to 0 such that $K \subseteq \overline{\mathrm{co}}(\{ x_n : n \in \mathbb{N} \})$, and there is a compact subset S of X that includes $\overline{\mathrm{co}}(\{ x_n : n \in \mathbb{N} \})$ such that (x_n) converges to 0 in the Banach space Y constructed from S as in Lemma 3.4.28. Let $\|\cdot\|_X$ and $\|\cdot\|_Y$ be the norms of X and Y respectively. For the rest of this argument, references to Y as a normed space are to $(Y, \|\cdot\|_Y)$ rather than to Y as a subspace of X. Since the "identity" map from Y into X is compact, there is a finite-rank member $\Phi_{K,\epsilon}$ of $B(Y, X)$ such that $\|\Phi_{K,\epsilon}x - x\|_X < \epsilon/2$ whenever $x \in K$. The goal is to find a finite-rank member of $B(X)$ that uniformly approximates the identity operator of X within ϵ on K, so it may be assumed that $\Phi_{K,\epsilon} \neq 0$. It follows that there are members $y_1^*, \dots, y_m^*$ of Y^* and a basis $z_1, \dots, z_m$ for $\Phi_{K,\epsilon}(Y)$ such that $\Phi_{K,\epsilon}y = \sum_{j=1}^m (y_j^* y) z_j$ whenever $y \in Y$. The proof that (a) $\Rightarrow$ (b) will be done once it is shown that if $y^* \in Y^*$ then there is an x^* in X^* such that $|y^*x - x^*x| < \epsilon/(2m \max\{ \|z_1\|_X, \dots, \|z_m\|_X \})$ whenever $x \in K$, for with elements $x_1^*, \dots, x_m^*$ of X^* near in this sense to $y_1^*, \dots, y_m^*$ respectively and $T_{K,\epsilon}x = \sum_{j=1}^m (x_j^* x) z_j$ whenever $x \in X$, it would follow that for each x in K,

$$\begin{aligned}
\|T_{K,\epsilon}x - x\|_X &\le \Big\| \sum_{j=1}^m (x_j^*x)z_j - \sum_{j=1}^m (y_j^*x)z_j \Big\|_X + \|\Phi_{K,\epsilon}x - x\|_X \\
&< \sum_{j=1}^m \big(|x_j^*x - y_j^*x| \, \|z_j\|_X \big) + \frac{\epsilon}{2} \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}$$

Suppose that $y^* \in Y^*$. Let $\delta = \epsilon/(2m \max\{ \|z_1\|_X, \dots, \|z_m\|_X \})$. Since $x_n \to 0$ in Y, there is an n_0 such that $|y^*x_n| < \delta/2$ when $n > n_0$. Let

$$K_{n_0} = 2\delta^{-1}\overline{\mathrm{co}}(\{ \alpha x_n : \alpha \in \mathbb{F},\ |\alpha| \le 1,\ n > n_0 \});$$

by Lemma 3.4.29 and the continuity of the map $x \mapsto x$ from Y into X, this is the same closed convex hull whether it is taken in Y or in X. Let

$$C = \{ y : y \in \langle x_1, \dots, x_{n_0} \rangle,\ \mathrm{Re}\, y^*y = 1 \}.$$

Then C is a closed subset of a finite-dimensional subspace of Y and so is closed in X, while K_{n_0} is, by Lemma 3.4.29, a compact subset of X. If $x_1, \ldots, x_{n_0} \in \ker(y^*)$, then by letting x^* be the zero element of X^* it would follow that $|y^*x - x^*x| = |y^*x| < \delta/2$ for every x in K, so it may be assumed that at least one of $y^*x_1, \ldots, y^*x_{n_0}$ is nonzero and therefore that $C \neq \varnothing$. Since $|y^*x| < 1$ when $x \in K_{n_0}$ and $|y^*x| \geq 1$ when $x \in C$, the sets K_{n_0} and C are disjoint, so by Theorem 2.2.28 there is an x^* in X^* such that

$$\max\{\operatorname{Re} x^*x : x \in K_{n_0}\} < \inf\{\operatorname{Re} x^*x : x \in C\}.$$

Notice that $\operatorname{Re} x^*$ must actually be constant on C, for otherwise $\operatorname{Re} x^*(C)$ would be $\mathbb{R}$. Since $0 \in \operatorname{Re} x^*(K_{n_0})$, it may be assumed that $\operatorname{Re} x^*x = 1$ whenever $x \in C$. It follows that $x^*x = y^*x$ whenever $x \in \langle x_1, \ldots, x_{n_0}\rangle$ and therefore that $x^*x_n = y^*x_n$ when $n = 1, \ldots, n_0$. Since K_{n_0} is balanced, it must be that $|x^*x_0| < \inf\{\operatorname{Re} x^*x : x \in C\} = 1$ whenever $x_0 \in K_{n_0}$, which implies that $|x^*x_n| < \delta/2$ when $n > n_0$. If $x \in K$, then let (t_n) be a sequence of nonnegative reals such that $\sum_n t_n \leq 1$ and $x = \sum_n t_n x_n$ and observe that

$$|x^*x - y^*x| = \left|\sum_n t_n(x^*x_n - y^*x_n)\right| \leq \sum_{n=n_0+1}^{\infty} t_n\big(|x^*x_n| + |y^*x_n|\big) < \delta.$$

This finishes the proof that (a) $\Rightarrow$ (b).

Now suppose that (b) holds. Let Z be a Banach space and let T be a compact operator from Z into X. By (b), there is a sequence (Ψ_n) of finite-rank members of $B(X)$ such that $\|\Psi_n x - x\|_X < 1/n$ for each x in $T(B_Z)$ and each n in $\mathbb{N}$. It follows that

$$\|\Psi_n T - T\| = \sup\{\|\Psi_n x - x\|_X : x \in T(B_Z)\} \leq \frac{1}{n}$$

for each n. Therefore T is the limit in $K(Z, X)$ of the sequence $(\Psi_n T)$ of finite-rank operators, so X has the approximation property. ■

The proof of the preceding theorem is essentially the one found in [156] with minor modifications to allow for the possibility of complex scalars.

For a long time one of the major open problems in Banach space theory, the *approximation problem,* was whether every Banach space has the approximation property. This was finally settled in the negative in a 1973 paper by Per Enflo [76], who found a separable reflexive Banach space, necessarily infinite-dimensional, that lacks the approximation property.[1]

---

[1]It had been established by Grothendieck in [96] that an example of a Banach space lacking the approximation property would also provide a negative answer to a question of Mazur from real analysis. On November 6, 1936, Mazur had entered his question

A. M. Davie [45] has published a simplified version of Enflo's example that is reproduced in [156].

See Grothendieck's memoir [96] for other properties equivalent to the approximation property. Another excellent source for more on the approximation property and its consequences is [156]. The approximation property will be briefly revisited in Section 4.1.

This section ends with a quick look at Hilbert's original version of the definition of a compact operator, which is essentially the following one.

3.4.33 Definition. (D. Hilbert, 1906 [103]). Suppose that X and Y are Banach spaces. A linear operator T from X into Y is *completely continuous* or a *Dunford-Pettis operator* if $T(K)$ is a compact subset of Y whenever K is a weakly compact subset of X.

Complete continuity is not really equivalent to compactness. As a property for linear operators between Banach spaces, complete continuity actually lies properly between compactness and boundedness, as the next two results and Exercise 3.51 together show.

3.4.34 Proposition. *Every compact linear operator from a Banach space into a Banach space is completely continuous.*

PROOF. Suppose that X and Y are Banach spaces, that $T \in K(X,Y)$, and that K is a weakly compact, hence bounded, subset of X. Then $T(K)$ is relatively compact, and is also weakly compact since T is weak-to-weak continuous. It follows easily that $T(K)$ is compact. ■

3.4.35 Proposition. *Every completely continuous linear operator from a Banach space into a Banach space is bounded.*

PROOF. Suppose that T is an unbounded linear operator from a Banach space X into a Banach space. Then there is a sequence (x_n) in B_X such that $\|Tx_n\| \geq n^2$ for each n. It follows that $\{ n^{-1}x_n : n \in \mathbb{N} \} \cup \{0\}$ is a weakly compact subset of X whose image under T is not bounded, hence not compact. ■

The following characterization of complete continuity is what Hilbert actually used as the definition of the property.

3.4.36 Proposition. *A linear operator from a Banach space into a Banach space is completely continuous if and only if it is weak-to-norm sequentially continuous.*

---

in the famous "Scottish book" of open problems kept at the Scottish Coffee House in Lwów, Poland, by Banach, Mazur, Stanisław Ulam, and other mathematicians in their circle; see [160, problem #153]. Mazur offered a live goose as the prize for a solution. About a year after solving the problem, Enflo traveled to Warsaw to give a lecture on his solution, after which he was awarded the goose.

PROOF. The forward implication is an easy consequence of the weak-to-weak continuity of completely continuous operators along with the fact that a subset of a Banach space consisting of the terms and limit of a weakly convergent sequence is weakly compact. The converse follows directly from the fact that weakly compact subsets of a normed space are weakly sequentially compact. ∎

Since sequential continuity does not in general imply continuity, the preceding proposition stops a bit short of claiming that every completely continuous operator is weak-to-norm continuous, and in fact not all completely continuous operators are; see Exercise 3.51. It does of course follow from the preceding proposition, and in fact directly from the definition of complete continuity, that every weak-to-norm continuous linear operator from a Banach space into a Banach space is completely continuous.

The reason that Hilbert is given joint credit with Riesz for founding the field of compact operator theory is that Hilbert was interested in linear operators with domain ℓ_2 when he gave the definition of complete continuity, and for such operators compactness and complete continuity are equivalent.

3.4.37 Theorem. *A linear operator from a reflexive Banach space into a (possibly nonreflexive) Banach space is compact if and only if it is completely continuous.*

PROOF. It is enough to prove that every completely continuous linear operator from a reflexive Banach space into a Banach space is compact, but this follows immediately from the weak compactness of the closed unit ball of the domain space. ∎

Though this section has concentrated on compact operators between arbitrary Banach spaces, the theory of compact operators between Hilbert spaces, especially the theory of the Banach algebra $K(H)$ where H is a Hilbert space, is particularly rich. In this situation, many of the results of this section can be proved in somewhat different and often simpler ways. The interested reader should have no difficulty finding sources in which this more specialized theory is developed. See, for example, Conway's text [43].

Exercises

3.39 Characterize the compact projections in a Banach space among all projections in that Banach space.

3.40 (a) Suppose that X is a reflexive Banach space. Prove that every member of $B(X, \ell_1)$ is compact.

(b) Suppose that Y is a reflexive Banach space. Prove that every member of $B(c_0, Y)$ is compact.

3.41 Give an example of a closed subalgebra of a Banach algebra that is not an ideal in the algebra.

3.42 Suppose that K is a compact Hausdorff space and that $S \subseteq C(K)$. Then S is *pointwise bounded* if $\{ f(x) : f \in S \}$ is a bounded subset of $\mathbb{F}$ whenever $x \in K$. Prove that if S is equicontinuous, then S is pointwise bounded if and only if S is a bounded subset of $C(K)$. (Thus, the boundedness condition in the statement of the Arzelà-Ascoli theorem can be replaced by pointwise boundedness, and in fact the theorem is often stated that way.)

3.43 Suppose that $g\colon [0,1] \times [0,1] \to \mathbb{F}$ is continuous. Define $T\colon C[0,1] \to C[0,1]$ by the formula $\big(T(f)\big)(t) = \int_0^1 g(t,s)f(s)\,ds$. Prove that T is a compact linear operator.

3.44 (a) Suppose that X and Y are infinite-dimensional normed spaces and that S is a one-to-one linear operator from Y^* into X^*. Prove that S is not weak*-to-norm continuous.

(b) Give an example of a compact linear operator T from a Banach space into a Banach space such that T^* is not weak*-to-norm continuous.

3.45 Prove the following improvement of Riesz's lemma for reflexive spaces: *If M is a proper closed subspace of a reflexive normed space X, then there is an x in S_X such that $d(x, M) = 1$.*

3.46 Lemma 3.4.21 is in a sense a generalization of the fact that the row rank and column rank of a matrix are the same. (At least, it is a generalization of that fact for square matrices.) Explain.

3.47 Suppose that X is a Banach space, that $T \in K(X)$, and that α is a nonzero scalar. Prove the following facts that are sometimes included in the Fredholm alternative.

(a) If $y \in X$, then $(\alpha I - T)(x) = y$ for some x in X if and only if $x^*y = 0$ whenever $x^* \in X^*$ and $(\alpha I^* - T^*)(x^*) = 0$.

(b) If $y^* \in X^*$, then $(\alpha I^* - T^*)(x^*) = y^*$ for some x^* in X^* if and only if $y^*x = 0$ whenever $x \in X$ and $(\alpha I - T)(x) = 0$.

(c) If $y \in X$ and the equation $(\alpha I - T)(x) = y$ has a solution, then the general solution of the equation is found by adding a particular solution of that equation to the general solution of the equation $(\alpha I - T)(x) = 0$.

(d) If $y^* \in X^*$ and the equation $(\alpha I^* - T^*)(x^*) = y^*$ has a solution, then the general solution of the equation is found by adding a particular solution of that equation to the general solution of the equation $(\alpha I^* - T^*)(x^*) = 0$.

3.48 Prove the following converse of Proposition 3.4.7: *If X is a separable Banach space, then there is some Banach space Y and a compact linear operator T from Y into X such that $T(Y)$ is dense in X.*

3.49 Prove that the collection of completely continuous linear operators from a Banach space X into a Banach space Y is a closed subspace of $B(X, Y)$.

3.50 Suppose that T is a linear operator from a normed space X into a normed space Y. Prove that the following are equivalent.

(a) The operator T is continuous.

(b) The set $T(K)$ is a compact subset of Y whenever K is a compact subset of X.

(c) The set $T(K)$ is a weakly compact subset of Y whenever K is a weakly compact subset of X.

Notice the analogy between this result and the definition of complete continuity.

3.51 (a) Suppose that Y is a Banach space. Prove that every member of $B(\ell_1, Y)$ is completely continuous. (Compare Exercise 3.40 (a).)

(b) Give an example of a linear operator from a nonreflexive Banach space into a reflexive Banach space that is completely continuous but not compact.

(c) Give an example of a linear operator from a Banach space into a Banach space that is bounded but not completely continuous.

(d) Give an example of a linear operator from a Banach space into a Banach space that is completely continuous but not weak-to-norm continuous.

3.5 Weakly Compact Operators

Since compact operators possess so many interesting properties, it is natural to ask how much of what has been shown in the preceding section survives if the linear operators being studied are only required to have the following weakened version of the compactness property.

3.5.1 Definition. (S. Kakutani, 1938 [125]; K. Yosida, 1938 [245]). Suppose that X and Y are Banach spaces. A linear operator T from X into Y is *weakly compact* if $T(B)$ is a relatively weakly compact subset of Y whenever B is a bounded subset of X. The collection of all weakly compact linear operators from X into Y is denoted by $K^w(X,Y)$, or by just $K^w(X)$ if $X = Y$.

The development of the theory of weakly compact operators in this section parallels that done for compact operators in the preceding section, up to but not including spectral theory. The reason for not saying anything here about the spectral theory of weakly compact operators is mentioned near the end of this section.

The first two results of this section place weak compactness between compactness and boundedness as a property for linear operators between Banach spaces. The first is obvious, while the second is an immediate consequence of the fact that every relatively weakly compact subset of a Banach space is bounded.

3.5.2 Proposition. *Every compact linear operator from a Banach space into a Banach space is weakly compact.*

3.5.3 Proposition. *Every weakly compact linear operator from a Banach space into a Banach space is bounded.*

Weak compactness is genuinely different from the other two properties for linear operators mentioned in the preceding two propositions. For example, it is easy to see that the identity operator on ℓ_1 is bounded but not weakly compact, while the identity operator on ℓ_2 is weakly compact but not compact.

Since a subset of a finite-dimensional subspace of a normed space is relatively weakly compact if and only if it is relatively compact, a finite-rank linear operator between Banach spaces is weakly compact if and only if it is compact. Thus, the most obvious analog for weak compactness of Proposition 3.4.3, that a finite-rank linear operator from a Banach space into a Banach space is weakly compact if and only if it is bounded, is trivially true. A more interesting analog can be obtained by first observing that, as an immediate consequence of Proposition 3.4.3, a bounded linear operator from a Banach space X into a Banach space Y is compact if either X or Y is finite-dimensional, and then recalling that weakly compact sets often play the same role in reflexive Banach spaces that compact sets do in finite-dimensional ones.

3.5.4 Proposition. *If X and Y are Banach spaces and either X or Y is reflexive, then every bounded linear operator from X into Y is weakly compact.*

PROOF. This follows easily from the relative weak compactness of bounded subsets of reflexive spaces and the weak-to-weak continuity of bounded linear operators between normed spaces. ■

The equivalence of the following characterizations of weak compactness for linear operators is easily proved using elementary arguments and the Eberlein-Šmulian theorem.

3.5.5 Proposition. *Suppose that T is a linear operator from a Banach space X into a Banach space Y. Then the following are equivalent.*

(a) *The operator T is weakly compact.*
(b) *The set $T(B_X)$ is a relatively weakly compact subset of Y.*
(c) *Every bounded sequence (x_n) in X has a subsequence (x_{n_j}) such that the sequence (Tx_{n_j}) converges weakly.*

The fact that the range of a compact operator is closed if and only if the operator has finite rank does have an analog for weakly compact operators, and again that analog is suggested by the similarity between the role of compact sets in finite-dimensional Banach spaces and that of weakly compact sets in reflexive Banach spaces.

3.5.6 Proposition. *The range of a weakly compact linear operator from a Banach space into a Banach space is closed if and only if the range of the operator is reflexive.*

PROOF. One implication is a trivial consequence of the fact that every reflexive subspace of a Banach space is closed. For the other, suppose that a weakly compact linear operator T from a Banach space X into a Banach space Y has closed range. Then T is an open mapping from X onto $T(X)$, so $T(B_X)$ is a relatively weakly compact subset of $T(X)$ that includes a neighborhood of 0 in $T(X)$. This implies that $B_{T(X)}$ is weakly compact and therefore that $T(X)$ is reflexive. ■

If T is a weakly compact linear operator from a Banach space X into a Banach space Y such that $T(X)$ is not known to be closed in Y, then it is not necessarily the case that $\overline{T(X)}$ is reflexive; see Exercise 3.52. However, it does follow that $\overline{T(X)}$ is the closed linear hull of the weakly compact set $\overline{T(B_X)}$, which yields the following analog of the fact that every compact linear operator from a Banach space into a Banach space has a separable (that is, compactly generated) range.

3.5.7 Proposition. *Suppose that T is a weakly compact linear operator from a Banach space X into a Banach space Y. Then $\overline{T(X)}$ is weakly compactly generated.*

The collections $K^w(X,Y)$ and $K^w(X)$ of Definition 3.5.1 have the same algebraic and topological closure properties that were shown to hold for $K(X,Y)$ and $K(X)$ in Section 3.4. Before showing this, it will be useful to have the following result relating the weak compactness of an operator to the location of the range of its second adjoint. This theorem is due to Gantmacher when the spaces are separable, while the general case is due to Nakamura.

3.5.8 Theorem. (V. Gantmacher, 1940 [83]; M. Nakamura, 1951 [173]). *Suppose that X and Y are Banach spaces, that $T \in B(X,Y)$, and that Q_Y is the natural map from Y into Y^{**}. Then T is weakly compact if and only if $T^{**}(X^{**}) \subseteq Q_Y(Y)$.*

PROOF. Let Q_X be the natural map from X into X^{**}. Suppose first that $T^{**}(X^{**}) \subseteq Q_Y(Y)$. By Proposition 2.6.24, the map Q_Y^{-1} is relative-weak*-to-weak continuous from $Q(Y)$ onto Y, and T^{**} is weak*-to-weak* continuous, so $Q_Y^{-1}T^{**}$ is weak*-to-weak continuous. It follows from the weak* compactness of $B_{X^{**}}$ that $Q_Y^{-1}T^{**}(B_{X^{**}})$ is weakly compact and therefore that its subset $Q_Y^{-1}T^{**}Q_X(B_X)$ is relatively weakly compact. Since

$Q_Y^{-1}T^{**}Q_X = T$ by Proposition 3.1.13, the set $T(B_X)$ is relatively weakly compact, so T is a weakly compact operator.

Suppose conversely that T is weakly compact. Then $\overline{Q_Y T(B_X)}^{w}$ is a weakly compact subset of Y^{**} and therefore is weakly* compact. This implies that

$$\overline{Q_Y T(B_X)}^{w^*} = \overline{Q_Y T(B_X)}^{w} = \overline{Q_Y T(B_X)},$$

where the last equality follows from the convexity of $Q_Y T(B_X)$ by Theorem 2.5.16. It is therefore a consequence of the weak*-to-weak* continuity of T^{**}, the Banach-Alaoglu theorem, Goldstine's theorem, and the fact that $Q_Y^{-1}T^{**}Q_X = T$, that

$$T^{**}(B_{X^{**}}) = \overline{T^{**}Q_X(B_X)}^{w^*} = \overline{Q_Y T(B_X)}^{w^*} = \overline{Q_Y T(B_X)} \subseteq Q_Y(Y),$$

from which it follows that $T^{**}(X^{**}) \subseteq Q_Y(Y)$. ■

3.5.9 Proposition. *Suppose that X and Y are Banach spaces, that R and S are weakly compact linear operators from X into Y, that $\alpha \in \mathbb{F}$, and that (T_n) is a sequence of weakly compact linear operators from X into Y that converges to some T in $B(X,Y)$. Then $R+S$, αR, and T are all weakly compact.*

PROOF. The claims about $R+S$ and αR follow easily from the equivalence of (a) and (c) in Proposition 3.5.5. For the weak compactness of T, just notice that if Q_Y is the natural map from Y into Y^{**}, then $T_n^{**}(X^{**})$ lies in the closed set $Q_Y(Y)$ for each n, so $T^{**}(X^{**}) \subseteq Q_Y(Y)$ since $T_n^{**} \to T^{**}$. ■

With X and Y as in the preceding proposition, the set $K^w(X,Y)$ is nonempty since it contains the zero operator from X into Y, so the proposition can be restated as follows.

3.5.10 Corollary. *If X and Y are Banach spaces, then $K^w(X,Y)$ is a closed subspace of $B(X,Y)$.*

The weak-to-weak continuity of a bounded linear operator from a Banach space into a Banach space assures that the operator maps relatively weakly compact sets to relatively weakly compact sets, and by definition such an operator maps bounded sets to bounded sets. The next result is an immediate consequence of this.

3.5.11 Proposition. *Suppose that X, Y, and Z are Banach spaces, that $S \in B(X,Y)$, and that $T \in B(Y,Z)$. If either S or T is weakly compact, then TS is weakly compact.*

3.5.12 Corollary. *If X is a Banach space, then $K^w(X)$ is a closed ideal in $B(X)$.*

As was done for compact operators in the preceding section, the weakly compact linear operators from a Banach space X into a Banach space Y will now be characterized among all members of $B(X, Y)$ in two different ways based on the behavior of their adjoints. The first characterization is just Schauder's theorem with compactness replaced by weak compactness, while the second, an analog of Theorem 3.4.16, involves a continuity property for the adjoint.

3.5.13 Gantmacher's Theorem. (V. Gantmacher, 1940 [83]). *A bounded linear operator from a Banach space into a Banach space is weakly compact if and only if its adjoint is weakly compact.*

PROOF. Let T be a bounded linear operator from a Banach space X into a Banach space Y. As usual, let the natural map from a Banach space Z into its second dual be denoted by Q_Z. Suppose first that T is weakly compact. Let y^{***} be an element of Y^{***}. To show that T^* is weakly compact, it is enough to show that $T^{***}y^{***} \in Q_{X^*}(X^*)$, for which it is enough to show that $T^{***}y^{***}$ is weakly* continuous on X^{**}. Suppose that (x_α^{**}) is a net in X^{**} that is weakly* convergent to some x^{**} in X^{**}. Then $T^{**}x_\alpha^{**} \xrightarrow{w^*} T^{**}x^{**}$. Since $T^{**}(X^{**}) \subseteq Q_Y(Y)$ and the relative weak and relative weak* topologies of $Q_Y(Y)$ as a subspace of Y^{**} are the same, it follows that $T^{**}x_\alpha^{**} \xrightarrow{w} T^{**}x^{**}$, and so

$$\langle x_\alpha^{**}, T^{***}y^{***}\rangle = \langle T^{**}x_\alpha^{**}, y^{***}\rangle \to \langle T^{**}x^{**}, y^{***}\rangle = \langle x^{**}, T^{***}y^{***}\rangle,$$

as required.

Suppose conversely that T^* is weakly compact. Then T^{**} is weakly compact by what has already been proved. Since $T = Q_Y^{-1}T^{**}Q_X$, it follows that T is weakly compact. ■

As with Theorem 3.5.8, Gantmacher established the following result for separable spaces while Nakamura obtained the general result.

3.5.14 Theorem. (V. Gantmacher, 1940 [83]; M. Nakamura, 1951 [173]). *A bounded linear operator from a Banach space into a Banach space is weakly compact if and only if its adjoint is weak*-to-weak continuous.*

PROOF. Let T be a bounded linear operator from a Banach space X into a Banach space Y. Suppose first that T is weakly compact, that (y_α^*) is a net in Y^* that is weakly* convergent to some y^*, and that $x^{**} \in X^{**}$. To show that T^* is weak*-to-weak continuous, it is enough to show that $x^{**}T^*y_\alpha^* \to x^{**}T^*y^*$, that is, that $\langle y_\alpha^*, T^{**}x^{**}\rangle \to \langle y^*, T^{**}x^{**}\rangle$. This follows from the fact that $T^{**}x^{**}$ is in the natural image of Y in Y^{**} and therefore is weakly* continuous.

Suppose conversely that T^* is weak*-to-weak continuous. Since each bounded subset of Y^* is relatively weakly* compact, it follows that T^*

maps bounded subsets of Y^* to relatively weakly compact subsets of X^* and is therefore weakly compact, so T is weakly compact. ■

As properties for linear operators, both weak compactness and complete continuity lie between compactness and boundedness, which suggests that the two properties might be related. In general, neither implies the other; see Exercise 3.58. However, it does happen that some common Banach spaces have the property that every weakly compact linear operator whose domain is that space is completely continuous. This property has a name given to it by Grothendieck in honor of N. Dunford and B. J. Pettis, who proved in a 1940 paper [66] that $L_1(S, \Sigma, \lambda)$ has the property when λ is Lebesgue measure on the σ-algebra Σ of Lebesgue-measurable subsets of a finite or infinite interval S in $\mathbb{R}^n$, where n is a positive integer.

3.5.15 Definition. (A. Grothendieck, 1953 [95]). A Banach space X has the *Dunford-Pettis property* if, for every Banach space Y, each weakly compact linear operator from X into Y is completely continuous.

It is obvious that every finite-dimensional Banach space has the Dunford-Pettis property. So do some infinite-dimensional ones.

3.5.16 Example. It can be shown that for every Banach space Y, each member of $B(\ell_1, Y)$ is completely continuous; see Exercise 3.51. Therefore ℓ_1 trivially has the Dunford-Pettis property. Another (perhaps not entirely independent) proof that ℓ_1 has this property can be found in Exercise 3.59.

However, not every infinite-dimensional Banach space has the Dunford-Pettis property. If X is an infinite-dimensional reflexive Banach space and I is the identity operator on X, then I is weakly compact but certainly not completely continuous since it does not map the closed unit ball of the space to a compact set. The next result follows from this.

3.5.17 Proposition. *No infinite-dimensional reflexive Banach space has the Dunford-Pettis property.*

The definition of the Dunford-Pettis property given above has the drawback that it involves every Banach space Y rather than just the space X in question. The following theorem gives characterizations of the Dunford-Pettis property that lack this defect.

3.5.18 Theorem. *Suppose that X is a Banach space. Then the following are equivalent.*

(a) *The space X has the Dunford-Pettis property.*
(b) *Every weakly compact linear operator from X into c_0 is completely continuous.*

(c) *For every sequence* (x_n) *in* X *converging weakly to some* x *and every sequence* (x_n^*) *in* X^* *converging weakly to some* x^*, *the sequence* $(x_n^* x_n)$ *converges to* $x^* x$.

(d) *For every sequence* (x_n) *in* X *converging weakly to* 0 *and every sequence* (x_n^*) *in* X^* *converging weakly to* 0, *the sequence* $(x_n^* x_n)$ *converges to* 0.

PROOF. Suppose that (d) does not hold. Let (x_n) and (x_n^*) be sequences converging weakly to 0 in X and X^* respectively such that $|x_n^* x_n| \geq \epsilon$ for some positive ϵ and each n. Define $T\colon X \to c_0$ by the formula $Tx = (x_n^* x)$. Then $T \in B(X, c_0)$, but T is not completely continuous since $\|Tx_m\| \geq \epsilon$ for each m. The immediate goal is to show that T is weakly compact. To this end, let (w_m) be a bounded sequence in X. It is enough to show that (Tw_m) has a weakly convergent subsequence. Let Q_X be the natural map from X into X^{**} and let x^{**} be a weak* accumulation point of $(Q_X w_m)$. It follows that there is a subsequence (w_{m_j}) of (w_m) such that $|x_k^* w_{m_j} - x^{**} x_k^*| < j^{-1}$ when $k = 1, \dots, j$. Now $(x^{**} x_n^*) \in c_0$, so it is enough to show that $(Tw_{m_j}) \xrightarrow{w} (x^{**} x_n^*)$, that is, that $w\text{-}\lim_j (x_n^* w_{m_j}) = (x^{**} x_n^*)$. Fix a positive integer n_0. It is now enough to show that $x_{n_0}^* w_{m_j} \to x^{**} x_{n_0}^*$, but this is guaranteed by the construction of the sequence (w_{m_j}). Therefore the operator T is weakly compact though not completely continuous, so (b) does not hold. This shows that (b) $\Rightarrow$ (d).

Suppose next that (a) does not hold, that is, that there is a Banach space Y and a member T of $K^w(X, Y)$ that is not completely continuous. Let (z_n) be a sequence in X such that $z_n \xrightarrow{w} 0$ but $\|Tz_n\| \geq \delta$ for some positive δ and every n. For each n, let y_n^* be a member of S_{Y^*} such that $y_n^* T z_n = \|Tz_n\|$. The adjoint T^* of T is weakly compact by Gantmacher's theorem, so the sequence $(T^* y_n^*)$ has a weakly convergent subsequence; it may be assumed that $T^* y_n^* \xrightarrow{w} z^*$ for some z^* in X^*. Then $(T^* y_n^*)(z_n) = y_n^* T z_n \geq \delta$ for each n, so $\big((T^* y_n^*)(z_n)\big)$ does not converge to $z^* 0$ (that is, to 0) even though $T^* y_n^* \xrightarrow{w} z^*$ and $z_n \xrightarrow{w} 0$. Therefore (c) fails, so (c) $\Rightarrow$ (a).

Now suppose that (d) holds. Let (v_n) be a sequence in X converging weakly to some v and let (v_n^*) be a sequence in X^* converging weakly to some v^*. Then

$$\begin{aligned} v_n^* v_n - v^* v &= (v_n^* - v^*)(v_n - v) + v_n^* v + v^* v_n - 2v^* v \\ &\to 0 + v^* v + v^* v - 2v^* v = 0, \end{aligned}$$

which shows that (d) $\Rightarrow$ (c). It is obvious that (a) $\Rightarrow$ (b), so the theorem is proved. ■

The spaces $L_1(\Omega, \Sigma, \mu)$ such that (Ω, Σ, μ) is a finite measure space have the Dunford-Pettis property, a fact that is essentially due to Dunford and Pettis [66], and the spaces $C(K)$ such that K is a compact Hausdorff space do also, as was shown by Grothendieck in a 1953 paper [95] and independently by Bartle, Dunford, and Schwartz in a 1955 paper [20]. These facts are proved in Diestel and Uhl's book [59] as Corollaries III.2.14 and VI.2.6, and an interesting discussion of the Dunford-Pettis property and its history appears on pp. 176–178 of the same work. See also Diestel's book [58], in which the fact that $C(K)$ and $L_1[0,1]$ have the Dunford-Pettis property is developed in exercises.

It was mentioned earlier that the spectral theory of weakly compact operators would not be developed in this section. The reason for this omission is illustrated by the following example.

3.5.19 Example. Let X be complex ℓ_2 and let $\mathbb{D}$ be the closed unit disc in $\mathbb{C}$. Let T_l and T_r be, respectively, the left-shift and right-shift operators on X, that is, the bounded linear operators from X into X defined by the formulas $T_l(\alpha_1, \alpha_2, \dots) = (\alpha_2, \alpha_3, \dots)$ and $T_r(\alpha_1, \alpha_2, \dots) = (0, \alpha_1, \alpha_2, \dots)$. It can be shown that $\sigma(T_l) = \mathbb{D}$ and $\sigma_p(T_l) = \mathbb{D}^\circ$, and that $\sigma(T_r) = \mathbb{D}$ and $\sigma_p(T_r) = \varnothing$; see Exercise 3.38. Also, the identity operator I on X has 1 as an eigenvalue, and the associated eigenspace, which is all of X, is certainly not finite-dimensional. By Proposition 3.5.4, the operators T_l, T_r, and I are all weakly compact. Together, they show that none of the conclusions of Theorem 3.4.23, or even reasonable generalizations of those conclusions such as the possibility that either the point spectrum or its complement in the spectrum must be countable, need hold for weakly compact operators from an infinite-dimensional Banach space into itself (except for the statement about the compactness of the spectrum, and that is true for the spectrum of every element of every Banach algebra with identity).

It does sometimes happen that an infinite-dimensional Banach space X has the property that the square of every member of $K^w(X)$ is compact, so the conclusions of Theorem 3.4.23 then hold for members T of $K^w(X)$ provided that T is replaced by T^2 in those conclusions. See Corollaries VI.7.5 and VI.8.13 in [67] for particular examples of this phenomenon.

Exercises

3.52 Give an example of a bounded linear operator T from some reflexive Banach space X onto a dense subspace of some nonreflexive Banach space Y. Conclude that T is weakly compact but $\overline{T(X)}$ is not reflexive.

3.53 Suppose that X and Y are Banach spaces, that $T \in L(X,Y)$, that M is a closed subspace of X such that $M \subseteq \ker(T)$, that π is the quotient map from X onto X/M, and that S is the unique map, automatically linear, from X/M into Y such that $T = S \circ \pi$; see the commutative diagram in Theorem 1.7.13.

(a) Prove that T is compact if and only if S is compact.

(b) Prove that T is weakly compact if and only if S is weakly compact.

3.54 (a) Prove that a linear operator from a Banach space into ℓ_1 is weakly compact if and only if it is compact.

(b) Prove that a linear operator from c_0 into a Banach space is weakly compact if and only if it is compact.

(c) Prove that neither c_0 nor ℓ_1 has any complemented infinite-dimensional reflexive subspaces.

3.55 Prove that if a weakly compact linear operator from one Banach space onto another has a reflexive kernel, then the domain of the operator is reflexive.

3.56 Prove the following converse of Proposition 3.5.7: *If Y is a weakly compactly generated Banach space, then there is some Banach space X and a weakly compact linear operator T from X into Y such that $T(X)$ is dense in Y.*

3.57 Obtain from Gantmacher's theorem another proof that a Banach space is reflexive if and only if its dual space is reflexive.

3.58 (a) Give an example of a linear operator from a Banach space into a Banach space such that the operator is weakly compact but not completely continuous.

(b) Give an example of a linear operator from a Banach space into a Banach space such that the operator is completely continuous but not weakly compact. (Exercise 3.51 might help for this part.)

3.59 Prove that every Banach space having Schur's property also has the Dunford-Pettis property.

3.60 (a) Suppose that X is a Banach space. Prove that if X^* has the Dunford-Pettis property, then so does X.

(b) Conclude from (a) and Exercise 3.59 that c_0 has the Dunford-Pettis property.

3.61 Why is there no section on weakly* compact operators in this book?

4

Schauder Bases

Much of the theory of finite-dimensional normed spaces that has been presented in this book is ultimately based on Theorem 1.4.12, which says that every linear operator from a finite-dimensional normed space X into any normed space Y is bounded. A careful examination of the proof of that theorem shows that it essentially amounts to demonstrating that if $x_1, \ldots, x_n$ is a vector space basis for X, then each of the linear "coordinate functionals" $\alpha_1 x_1 + \cdots + \alpha_n x_n \mapsto \alpha_m$, $m = 1, \ldots, n$, is bounded; this can be seen from the nature of the norm $/\cdot/$ used in the proof and the argument near the end of the proof that there is no sequence (z_j) in B_X such that $/Iz_j/ \geq j$ for each j. It should not be too surprising that many topological results about finite-dimensional normed spaces are ultimately based on the continuity of the members of the family $\mathfrak{B}^\#$ of coordinate functionals for some basis $\mathfrak{B}$ for the space, since it is an easy consequence of Proposition 2.4.8, Theorem 2.4.11, and the uniqueness of Hausdorff vector topologies for finite-dimensional vector spaces that the norm topology of the space is the $\mathfrak{B}^\#$ topology of the space.

Though it is possible for an incomplete infinite-dimensional normed space to have a vector space basis $\mathfrak{B}$ such that all coordinate functionals for $\mathfrak{B}$ are bounded, this can never happen for an infinite-dimensional Banach space; see Exercises 4.1 and 4.2. Thus, the topology of an infinite-dimensional Banach space is not nearly so closely tied to coordinate functionals with respect to a vector space basis as is the topology of a finite-dimensional Banach space. It is primarily for this reason that vector space bases have not been used very much in this book's exploration of the general theory of Banach spaces.

Many of the classical infinite-dimensional Banach spaces do have sequences of elements with properties enough like those of a vector space basis to allow some arguments involving bases in finite-dimensional vector spaces to be carried over almost unchanged. For example, consider the sequence (e_n) of standard unit vectors of c_0. The set consisting of the terms of this sequence is not a vector space basis for c_0, nor is any other countable subset of c_0 by Theorem 1.5.8. However, this sequence does have the property that every element (α_n) of c_0 can be written in exactly one way as a "linear combination" of these unit vectors, namely, as $\sum_n \alpha_n e_n$, provided that *infinite* sums are allowed. Notice also that for each positive integer m, the "coordinate functional" $(\alpha_n) \mapsto \alpha_m$ is a bounded linear functional. Compare the proof of the separability of c_0 given in Example 1.12.6 to the proof of the separability of all finite-dimensional normed spaces in Example 1.12.8 to see one instance in which the standard unit vectors of c_0 are used to extend a finite-dimensional argument to that space.

The purpose of this chapter is to study sequences in Banach spaces such that every member of the space can be written in exactly one way as an "infinite linear combination" of the terms of the sequence, and to see what can be learned about the structure of Banach spaces having such *Schauder bases.*

4.1 First Properties of Schauder Bases

Notice that the following definition pertains to ordered sequences (x_n) rather than unordered sets of the form $\{ x_n : n \in \mathbb{N} \}$.

4.1.1 Definition. (J. Schauder, 1927 [206]). A sequence (x_n) in a Banach space X is a *Schauder basis* for X if for each x in X there is a unique sequence (α_n) of scalars such that $x = \sum_n \alpha_n x_n$.

Schauder included in his definition the requirement that each of the coordinate functionals $\sum_n \alpha_n x_n \mapsto \alpha_m$ such that $m \in \mathbb{N}$ be continuous. As will be shown in Corollary 4.1.16, that actually follows from the rest of the definition.

The following generalization of the notion of a Schauder basis also has its uses.

4.1.2 Definition. A sequence (x_n) in a Banach space is a *Schauder basic sequence* if it is a Schauder basis for $[\{ x_n : n \in \mathbb{N} \}]$.

Henceforth, whenever reference is made to a basis for a Banach space or a basic sequence in a Banach space, the reference is to a Schauder basis or a Schauder basic sequence unless stated otherwise. To avoid confusion, vector space bases are often called *Hamel bases* after the German analyst and applied mathematician Georg Karl Wilhelm Hamel (1877–1954).

4.1.3 Example. If X is c_0 or ℓ_p such that $1 \le p < \infty$, then it is easy to check that the sequence (e_n) of standard unit vectors of X is a basis for X and that $(\alpha_n) = \sum_n \alpha_n e_n$ whenever $(\alpha_n) \in X$. However, the sequence (e_n) is not a basis for ℓ_∞. For example, there is no sequence (α_n) of scalars such that $(1, 1, 1, \dots) = \sum_n \alpha_n e_n$. See Exercise 4.3.

Whenever the sequence (e_n) lies in the unit sphere of a Banach space of sequences of scalars and is a basis for the space, the sequence will be called the *standard unit vector basis* for the space.

The basis of the preceding example has the special property that each of its terms has norm 1. As will now be shown, every Banach space having any basis whatever has a basis with this property that can be obtained from the original basis in the obvious way.

4.1.4 Definition. A basic sequence (x_n) in a Banach space is *bounded* if $0 < \inf_n \|x_n\| \le \sup_n \|x_n\| < +\infty$, and is *normalized* if each x_n has norm 1.

The use of the terms "bounded basic sequence" and "bounded basis" in the sense of the preceding definition is common but unfortunate, since the usual meaning of the word "bounded" when used to describe a sequence (x_n) in a Banach space does not imply that $\inf_n \|x_n\| > 0$. To avoid confusion, the word will not be applied to basic sequences in this book.

4.1.5 Proposition. *Suppose that (x_n) is a basis for a Banach space X and that (λ_n) is a sequence of nonzero scalars. Then $(\lambda_n x_n)$ is also a basis for X.*

PROOF. It is easy to check that for each member x of X there is a unique sequence of scalars (α_n) such that $x = \sum_n \alpha_n \lambda_n x_n$, as required. ■

The uniqueness of basis expansions implies that the terms of a basis are nonzero, so the following result is an immediate consequence of the preceding proposition.

4.1.6 Corollary. *If (x_n) is a basis for a Banach space, then $\left(\|x_n\|^{-1} x_n\right)$ is a normalized basis for the space.*

Just as each of the classical separable Banach sequence spaces of Example 4.1.3 has the sequence of standard unit vectors as a basis, every Banach space with a basis can be viewed in a natural way as a sequence space for which the sequence (e_n) of standard coordinate vectors is a basis (though it would be misleading to call this basis the standard *unit* vector basis for the space, since the terms of the sequence might not have norm 1). The following proposition shows how this is done, and is obvious enough that it is really just an observation.

4.1.7 Proposition. *Suppose that X is a Banach space with a basis (x_n). Let Y be the collection of all sequences (α_n) of scalars such that $\sum_n \alpha_n x_n$ converges, treated as a normed space with the usual vector space operations for a space of sequences of scalars and with the norm $\|\cdot\|_Y$ given by the formula $\|(\alpha_n)\|_Y = \|\sum_n \alpha_n x_n\|$. Then Y is a Banach space having the sequence (e_n) of standard coordinate vectors as a basis, and the map $\sum_n \alpha_n x_n \mapsto (\alpha_n)$ is an isometric isomorphism from X onto Y.*

The isometric isomorphism of the preceding proposition maps the given basis (x_n) for X onto a basis (e_n) for Y. This turns out to be a general property of Banach space isomorphisms.

4.1.8 Proposition. *Suppose that X and Y are Banach spaces, that T is an isomorphism from X into Y, and that (x_n) is a basic sequence in X. Then (Tx_n) is a basic sequence in Y. In particular, if (x_n) is a basis for X and T maps X onto Y, then (Tx_n) is a basis for Y.*

PROOF. Since T maps $[\{x_n : n \in \mathbb{N}\}]$ onto $[\{Tx_n : n \in \mathbb{N}\}]$, it may be assumed that (x_n) is a basis for X and that $T(X) = Y$. It is then enough to show that (Tx_n) is a basis for Y, which follows easily from the fact that $T(\sum_n \alpha_n x_n) = \sum_n \alpha_n Tx_n$ whenever $\sum_n \alpha_n x_n \in X$. ∎

The spaces with bases in Example 4.1.3 are infinite-dimensional and separable. In fact, every Banach space with a basis must be separable by Proposition 1.12.1 (a) since the space is the closed linear hull of the collection of basis elements. Also, every basic sequence is linearly independent since no element of the underlying Banach space can be written in two different ways as a finite linear combination of the terms of the sequence. Therefore, Banach spaces having bases must be infinite-dimensional. These observations are summarized formally in the following two results.

4.1.9 Proposition. *Every basic sequence in a Banach space is linearly independent.*

4.1.10 Proposition. *Every Banach space having a basis is infinite-dimensional and separable.*

It can be shown that every separable Banach space is isometrically isomorphic to a subspace of $C[0,1]$; see Exercise 2.74. It is therefore of some interest that $C[0,1]$ has a basis.

4.1.11 Example: *The classical Schauder basis for* $C[0,1]$. This example is from Schauder's 1927 paper [206] that introduced the notion of Schauder

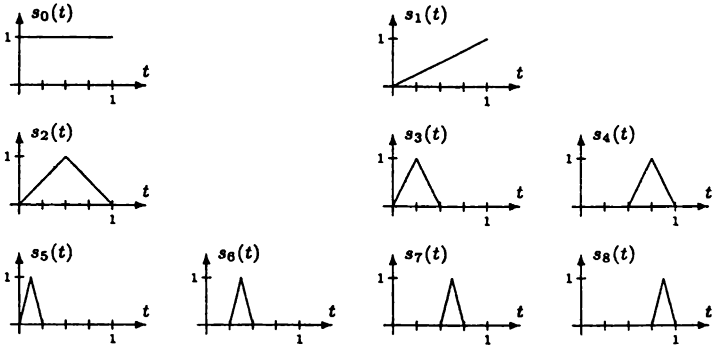


FIGURE 4.1. The first few terms of the classical Schauder basis for $C[0,1]$.

bases. Define the sequence $(s_n)_{n=0}^{\infty}$ of members of $C[0,1]$ as follows. Let $s_0(t) = 1$ and $s_1(t) = t$. When $n \geq 2$, define s_n by letting m be the positive integer such that $2^{m-1} < n \leq 2^m$, then let

$$s_n(t) = \begin{cases} 2^m\Big(t - \big(\frac{2n-2}{2^m} - 1\big)\Big) & \text{if } \frac{2n-2}{2^m} - 1 \leq t < \frac{2n-1}{2^m} - 1; \\ 1 - 2^m\Big(t - \big(\frac{2n-1}{2^m} - 1\big)\Big) & \text{if } \frac{2n-1}{2^m} - 1 \leq t < \frac{2n}{2^m} - 1; \\ 0 & \text{otherwise.} \end{cases}$$

See Figure 4.1 for the graphs of s_0 through s_8, which should demystify these formulas a bit.

Suppose that $f \in C[0,1]$. Define a sequence $(p_n)_{n=0}^{\infty}$ in $C[0,1]$ in the following way. Let

$$\begin{aligned} p_0 &= f(0)s_0, \\ p_1 &= p_0 + \big(f(1) - p_0(1)\big)s_1, \\ p_2 &= p_1 + \big(f(1/2) - p_1(1/2)\big)s_2, \\ p_3 &= p_2 + \big(f(1/4) - p_2(1/4)\big)s_3, \\ p_4 &= p_3 + \big(f(3/4) - p_3(3/4)\big)s_4, \\ p_5 &= p_4 + \big(f(1/8) - p_4(1/8)\big)s_5, \\ p_6 &= p_5 + \big(f(3/8) - p_5(3/8)\big)s_6, \\ p_7 &= p_6 + \big(f(5/8) - p_6(5/8)\big)s_7, \\ p_8 &= p_7 + \big(f(7/8) - p_7(7/8)\big)s_8, \end{aligned}$$

and so forth. Then p_0 is the constant function that agrees with f at 0, while p_1 agrees with f at 0 and 1 and interpolates linearly in between, and p_2 agrees with f at 0, 1, and 1/2 and interpolates linearly in between, and so forth. For each nonnegative integer n, let α_n be the coefficient of s_n in the formula for p_n. Then $p_m = \sum_{n=0}^{m} \alpha_n s_n$ for each m. It follows easily from the uniform continuity of f that $\lim_m \|p_m - f\|_\infty = 0$ and therefore that $f = \sum_{n=0}^\infty \alpha_n s_n$.

Now let $(\beta_n)_{n=0}^\infty$ be any sequence of scalars such that $f = \sum_{n=0}^\infty \beta_n s_n$. Then $\sum_{n=0}^\infty (\alpha_n - \beta_n)s_n = 0$, which implies that $\sum_{n=0}^\infty (\alpha_n - \beta_n)s_n(t) = 0$ when $t = 0, 1, 1/2, 1/4, 3/4, 1/8, 3/8, 5/8, 7/8, \dots$, from which it quickly follows that $\alpha_n = \beta_n$ for each n. Therefore there is a unique sequence $(\gamma_n)_{n=0}^\infty$ of scalars such that $f = \sum_{n=0}^\infty \gamma_n s_n$, and so the sequence $(s_n)_{n=0}^\infty$ is a basis for $C[0,1]$. Notice that this basis is normalized.

The next order of business is to prove the continuity of two types of natural linear maps on Banach spaces having bases.

4.1.12 Definition. Suppose that a Banach space X has a basis (x_n). For each positive integer m, the m^{th} *coordinate functional* x_m^* for (x_n) is the map $\sum_n \alpha_n x_n \mapsto \alpha_m$ from X into $\mathbb{F}$, and the m^{th} *natural projection* P_m for (x_n) is the map $\sum_n \alpha_n x_n \mapsto \sum_{n=1}^m \alpha_n x_n$ from X into X.

With all notation as in the preceding definition, it is clear that each x_m^* is a linear functional on X and that each P_m is a projection from X onto $\langle\{x_1, \dots, x_m\}\rangle$. To show that these maps are bounded, it is convenient to work not with the original norm of the underlying Banach space, but instead with the following one.

4.1.13 Definition. Suppose that (x_n) is a basis for a Banach space X. Then the (x_n) *norm* of X is defined by the formula $\|\sum_n \alpha_n x_n\|_{(x_n)} = \sup_m \|\sum_{n=1}^m \alpha_n x_n\|$.

It is important to notice that the norm $\|\cdot\|_{(x_n)}$ in the preceding definition depends not only on (x_n), but also on the space's original norm.

4.1.14 Theorem. *Suppose that (x_n) is a basis for a Banach space X. Then the (x_n) norm of X is a Banach norm equivalent to the original norm of X, and $\|x\|_{(x_n)} \geq \|x\|$ for each x in X.*

PROOF. Throughout this proof, symbols denoting convergent series represent series convergent with respect to the original norm of X, not the (x_n) norm, even when such series appear inside (x_n) norm function symbols. (Of course, one consequence of this theorem is that there is no difference between these two types of series convergence.)

Suppose that a member of X has expansion $\sum_n \alpha_n x_n$ in terms of the basis (x_n). Then

$$\left\|\sum_n \alpha_n x_n\right\|_{(x_n)} = \sup_m \left\|\sum_{n=1}^m \alpha_n x_n\right\| \geq \lim_m \left\|\sum_{n=1}^m \alpha_n x_n\right\| = \left\|\sum_n \alpha_n x_n\right\|,$$

which is the inequality claimed in the statement of the theorem.

It is easy to check that $\|\cdot\|_{(x_n)}$ really is a norm. To see that it is complete, let $(\sum_n \alpha_{n,j}x_n)_{j=1}^\infty$ be a sequence in X that is Cauchy with respect to $\|\cdot\|_{(x_n)}$. If $j_1, j_2, k \in \mathbb{N}$ and $k \geq 2$, then

$$\begin{aligned}|\alpha_{k,j_1} - \alpha_{k,j_2}|\,\|x_k\| &= \left\|\sum_{n=1}^{k}(\alpha_{n,j_1} - \alpha_{n,j_2})x_n - \sum_{n=1}^{k-1}(\alpha_{n,j_1} - \alpha_{n,j_2})x_n\right\| \\ &\leq 2\left\|\sum_n(\alpha_{n,j_1} - \alpha_{n,j_2})x_n\right\|_{(x_n)}\end{aligned}$$

and $|\alpha_{1,j_1} - \alpha_{1,j_2}|\,\|x_1\| \leq \|\sum_n(\alpha_{n,j_1} - \alpha_{n,j_2})x_n\|_{(x_n)}$, from which it follows that, for each n, the sequence $(\alpha_{n,j})_{j=1}^\infty$ is Cauchy and therefore convergent. Let (α_n) be the resulting sequence of limits. Fix a positive ϵ, and let j_ϵ be a positive integer such that if $j, j' \geq j_\epsilon$, then

$$\left\|\sum_{n=1}^{m} \alpha_{n,j'}x_n - \sum_{n=1}^{m} \alpha_{n,j}x_n\right\| \leq \left\|\sum_n \alpha_{n,j'}x_n - \sum_n \alpha_{n,j}x_n\right\|_{(x_n)} < \frac{\epsilon}{3}$$

for each m. Letting $j = j_\epsilon$ and letting j' tend to infinity shows that

$$\left\|\sum_{n=1}^{m} \alpha_n x_n - \sum_{n=1}^{m} \alpha_{n,j_\epsilon}x_n\right\| \leq \frac{\epsilon}{3} \tag{4.1}$$

for each m. It follows that if $m_2 \geq m_1 > 1$, then

$$\begin{aligned}&\left\|\sum_{n=m_1}^{m_2} \alpha_n x_n - \sum_{n=m_1}^{m_2} \alpha_{n,j_\epsilon}x_n\right\| \\ &\quad\leq \left\|\sum_{n=1}^{m_2} \alpha_n x_n - \sum_{n=1}^{m_2} \alpha_{n,j_\epsilon}x_n\right\| + \left\|\sum_{n=1}^{m_1-1} \alpha_n x_n - \sum_{n=1}^{m_1-1} \alpha_{n,j_\epsilon}x_n\right\| \\ &\quad\leq \frac{2\epsilon}{3}.\end{aligned}$$

Now let m_ϵ be such that if $m_2 \geq m_1 > m_\epsilon$, then $\|\sum_{n=m_1}^{m_2} \alpha_{n,j_\epsilon}x_n\| < \epsilon/3$. It follows that

$$\left\|\sum_{n=m_1}^{m_2} \alpha_n x_n\right\| \leq \left\|\sum_{n=m_1}^{m_2} \alpha_n x_n - \sum_{n=m_1}^{m_2} \alpha_{n,j_\epsilon}x_n\right\| + \left\|\sum_{n=m_1}^{m_2} \alpha_{n,j_\epsilon}x_n\right\| < \epsilon$$

when $m_2 \geq m_1 > m_\epsilon$, so $\sum_n \alpha_n x_n$ converges. Now (4.1) still holds if j_ϵ is replaced by any j such that $j \geq j_\epsilon$, so making that substitution and taking the supremum over all m shows that $\lim_j \|\sum_n \alpha_n x_n - \sum_n \alpha_{n,j}x_n\|_{(x_n)} = 0$, which finishes the proof that $\|\cdot\|_{(x_n)}$ is a Banach norm.

Finally, the inequality verified at the beginning of this proof implies that the identity operator from $(X, \|\cdot\|_{(x_n)})$ onto $(X, \|\cdot\|)$ is continuous, which by Corollary 1.6.8 assures that the two norms are equivalent. ∎

The continuity of coordinate functionals and natural projections follows easily from the preceding theorem.

4.1.15 Theorem. *Each natural projection associated with a basis for a Banach space is bounded.*

PROOF. Suppose that (x_n) is a basis for a Banach space X, that $m \in \mathbb{N}$, and that P_m is the m^{th} natural projection for (x_n). If $\sum_n \alpha_n x_n \in X$, then

$$\left\| P_m\Big(\sum_n \alpha_n x_n\Big)\right\|_{(x_n)} = \left\|\sum_{n=1}^m \alpha_n x_n\right\|_{(x_n)} \le \left\|\sum_n \alpha_n x_n\right\|_{(x_n)},$$

from which the continuity of P_m is immediate. ■

4.1.16 Corollary. (S. Banach, 1932 [13, p. 111]). *Each coordinate functional associated with a basis for a Banach space is bounded.*

PROOF. Let (x_n^*) and (P_n) be, respectively, the sequence of coordinate functionals and the sequence of natural projections for a basis (x_n) for a Banach space X, and let m be a positive integer. If $m \ge 2$, then x_m^* is the map $\sum_n \alpha_n x_n \mapsto (P_m - P_{m-1})(\sum_n \alpha_n x_n) = \alpha_m x_m \mapsto \alpha_m$, while x_1^* is the map $\sum_n \alpha_n x_n \mapsto P_1(\sum_n \alpha_n x_n) = \alpha_1 x_1 \mapsto \alpha_1$. In either case, the map x_m^* is clearly continuous. ■

4.1.17 Corollary. *If $\{ P_n : n \in \mathbb{N} \}$ is the collection of natural projections associated with a basis for a Banach space, then $\sup_n \|P_n\|$ is finite.*

PROOF. This is an easy consequence of the uniform boundedness principle, but also follows directly from the inequality in the proof of Theorem 4.1.15 by noting that, in the notation of that proof, if I is the identity map from $(X, \|\cdot\|_{(x_n)})$ onto $(X, \|\cdot\|)$, then

$$\begin{aligned}
\left\| P_m\Big(\sum_n \alpha_n x_n\Big)\right\| &\le \|I\| \left\| P_m\Big(\sum_n \alpha_n x_n\Big)\right\|_{(x_n)} \\
&\le \|I\| \left\|\sum_n \alpha_n x_n\right\|_{(x_n)} \\
&\le \|I\|\,\|I^{-1}\| \left\|\sum_n \alpha_n x_n\right\|
\end{aligned}$$

whenever $m \in \mathbb{N}$ and $\sum_n \alpha_n x_n \in X$. ■

4.1.18 Definition. Let $\{ P_n : n \in \mathbb{N} \}$ be the collection of natural projections associated with a basis for a Banach space. Then $\sup_n \|P_n\|$ is the ***basis constant*** for that basis. The basis is *monotone* or *orthogonal* if its basis constant is 1.

The term *monotone* for a basis with basis constant 1 was introduced by M. M. Day in his 1958 book [53], while the older term *orthogonal* for such bases dates back to independent studies of them by V. Ya. Kozlov [142] and R. C. James [110] published in 1950 and 1951 respectively.

Since each natural projection for a basis has norm at least 1, basis constants are always greater than or equal to 1, and a basis is monotone if and only if each of the natural projections for the basis has norm 1.

Of course, a basic sequence is a basis for the closed linear hull of the set consisting of the terms of the sequence and therefore has a basis constant associated with it. As one would expect, a monotone basic sequence is one whose basis constant is 1. In general, terminology used for bases automatically extends to basic sequences in this fashion as long as the extension makes sense.

It must be emphasized that basis constants depend on the norm of the space. In particular, it will follow from Exercise 4.11 and Corollary 4.1.22 that every Banach space with a basis has a nonmonotone basis (x_n) that becomes monotone when the space is renormed with $\|\cdot\|_{(x_n)}$.

4.1.19 Example. If X is c_0 or ℓ_p such that $1 \le p < \infty$, then the standard unit vector basis for X is monotone.

Some useful characterizations of basis constants and monotone bases are given in the following two propositions. It is from the equivalence of (a) and (c) in the second of these propositions that monotone bases get their name. (The reader interested in knowing why such bases were once called *orthogonal* should be able to reconstruct the reason from this same equivalence and the comments preceding Lemma 4.1.28.)

4.1.20 Proposition. *Suppose that (x_n) is a basis for a Banach space X and that K is the basis constant for (x_n). Then K is the smallest real number M such that*

$$\left\|\sum_{n=1}^{m} \alpha_n x_n\right\| \le M\left\|\sum_{n} \alpha_n x_n\right\|$$

whenever $\sum_n \alpha_n x_n \in X$ and $m \in \mathbb{N}$, which is in turn the smallest real number M such that

$$\left\|\sum_{n=1}^{m_1} \alpha_n x_n\right\| \le M\left\|\sum_{n=1}^{m_2} \alpha_n x_n\right\|$$

whenever $m_1, m_2 \in \mathbb{N}$, $m_1 \le m_2$, and $\alpha_1, \ldots, \alpha_{m_2} \in \mathbb{F}$.

PROOF. It is a straightforward consequence of Definition 4.1.18 that

$$K = \sup\left\{ \frac{\|\sum_{n=1}^{m} \alpha_n x_n\|}{\|\sum_n \alpha_n x_n\|} : \sum_n \alpha_n x_n \in X \setminus \{0\},\, m \in \mathbb{N} \right\},$$

from which the proposition readily follows. ■

The equivalence of (a) and (b) in the next result follows immediately from the preceding proposition. It is clear that (b) and (c) are equivalent, while the equivalence of (c) and (d) is an easy consequence of the definition of the (x_n) norm.

4.1.21 Proposition. *Suppose that (x_n) is a basis for a Banach space X. Then the following are equivalent.*

(a) *The basis (x_n) is monotone.*

(b) $\|\sum_{n=1}^m \alpha_n x_n\| \le \|\sum_n \alpha_n x_n\|$ *for each positive integer m and each member $\sum_n \alpha_n x_n$ of X.*

(c) $\|\sum_{n=1}^m \alpha_n x_n\| \le \|\sum_{n=1}^{m+1} \alpha_n x_n\|$ *for each positive integer m and each collection of $m+1$ scalars $\alpha_1, \dots, \alpha_{m+1}$.*

(d) *The original norm of X and the (x_n) norm of X are the same.*

4.1.22 Corollary. *If (x_n) is a basis for a Banach space X, then (x_n) is monotone with respect to the (x_n) norm of X.*

The following definition is suggested by the equivalence of (a) and (c) in the preceding proposition.

4.1.23 Definition. (M. G. Krein, M. A. Krasnoselskiĭ, and D. P. Milman, 1948 [144]). A basis (x_n) for a Banach space is *strictly monotone* if $\|\sum_{n=1}^m \alpha_n x_n\| < \|\sum_{n=1}^{m+1} \alpha_n x_n\|$ for each positive integer m and each collection of $m+1$ scalars $\alpha_1, \dots, \alpha_{m+1}$ such that $\alpha_{m+1} \ne 0$.

For example, the standard unit vector basis for each space ℓ_p such that $1 \le p < \infty$ is strictly monotone, while the corresponding basis for c_0 is monotone but not strictly monotone.

So far, the results of this section have been about conclusions that can be drawn when it is known that certain sequences are basic rather than properties sequences can have that assure that they are basic. The next theorem is probably the most important result of the latter kind.

4.1.24 Theorem. (S. Banach, 1932 [13]). *A sequence (x_n) in a Banach space X is a basis for X if and only if*

(1) *each x_n is nonzero;*

(2) *there is a real number M such that*

$$\left\|\sum_{n=1}^{m_1} \alpha_n x_n\right\| \le M \left\|\sum_{n=1}^{m_2} \alpha_n x_n\right\|$$

whenever $m_1, m_2 \in \mathbb{N}$, $m_1 \le m_2$, and $\alpha_1, \dots, \alpha_{m_2} \in \mathbb{F}$; and

(3) $[\{x_n : n \in \mathbb{N}\}] = X$.

PROOF. By what has already been proved in this section, the sequence (x_n) has properties (1), (2), and (3) if it is a basis for X.

For the proof of the converse, assume that (x_n) satisfies (1), (2), and (3). Suppose first that (β_n) and (γ_n) are sequences of scalars such that $\sum_n \beta_n x_n$ and $\sum_n \gamma_n x_n$ both converge and are equal. By (2),

$$|\beta_1 - \gamma_1| \, \|x_1\| \leq M \left\| \sum_n \beta_n x_n - \sum_n \gamma_n x_n \right\| = 0,$$

so $\beta_1 = \gamma_1$. It then follows by induction that $\beta_n = \gamma_n$ for each positive integer n. Thus, for no element x of X is there more than one sequence (α_n) of scalars such that $x = \sum_n \alpha_n x_n$.

For each positive integer m and each finitely nonzero sequence (α_n) of scalars, let $p_m(\sum_n \alpha_n x_n) = \sum_{n=1}^m \alpha_n x_n$. It follows from (2) that each p_m is a bounded linear operator from $\langle\{ x_n : n \in \mathbb{N} \}\rangle$ onto $\langle\{x_1, \ldots, x_m\}\rangle$ having norm no more than M, and so, by Theorem 1.9.1, has a bounded linear extension P_m from X onto $\langle\{x_1, \ldots, x_m\}\rangle$ with norm no more than M. For each x in X, each y in $\langle\{ x_n : n \in \mathbb{N} \}\rangle$, and each positive integer m,

$$\begin{aligned} \|P_m x - x\| &\leq \|P_m x - P_m y\| + \|P_m y - y\| + \|y - x\| \\ &\leq (M+1)\|x - y\| + \|P_m y - y\|. \end{aligned}$$

Letting m tend to infinity shows that $\limsup_m \|P_m x - x\| \leq (M+1)\|x-y\|$ whenever $x \in X$ and $y \in \langle\{ x_n : n \in \mathbb{N} \}\rangle$. Since y is an arbitrary member of a dense subset of X, it follows that $\limsup_m \|P_m x - x\| = 0$ whenever $x \in X$, that is, that $\lim_m P_m x = x$ for each x in X.

Fix an x in X and let α_1 be such that $P_1 x = \alpha_1 x_1$. Now $P_1 P_2 y = P_1 y$ whenever y is in the dense subset $\langle\{ x_n : n \in \mathbb{N} \}\rangle$ of X, from which it follows that $P_1 P_2 = P_1$. Therefore, there is an α_2 such that $P_2 x = \sum_{n=1}^2 \alpha_n x_n$. An easy induction argument produces a sequence (α_n) of scalars such that $P_m x = \sum_{n=1}^m \alpha_n x_n$ for each m, so $x = \lim_m P_m x = \sum_n \alpha_n x_n$. It follows from this and the uniqueness assertion proved earlier that (x_n) is a basis for X. ■

4.1.25 Corollary. *A sequence (x_n) in a Banach space is basic if and only if each x_n is nonzero and there is a real number M such that*

$$\left\| \sum_{n=1}^{m_1} \alpha_n x_n \right\| \leq M \left\| \sum_{n=1}^{m_2} \alpha_n x_n \right\|$$

whenever $m_1, m_2 \in \mathbb{N}$, $m_1 \leq m_2$, and $\alpha_1, \ldots, \alpha_{m_2} \in \mathbb{F}$.

4.1.26 Corollary. *Every subsequence of a basic sequence in a Banach space is itself a basic sequence.*

4.1.27 Example: *The Haar basis for $L_p[0,1]$, $1 \leq p < \infty$.* Suppose that $1 \leq p < \infty$. Define a sequence (h_n) in $L_p[0,1]$ as follows. Let h_1 be 1

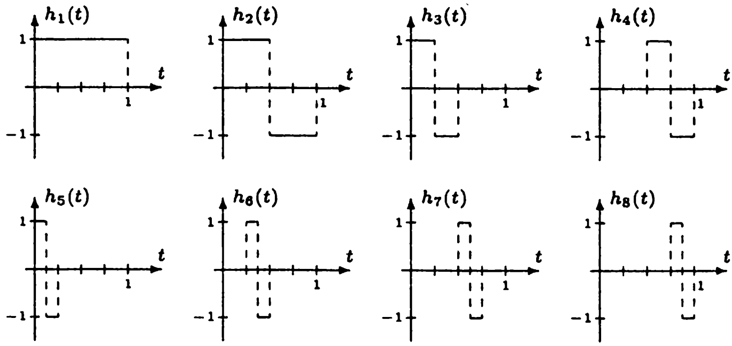


FIGURE 4.2. The first few terms of the Haar basis for $L_p[0,1]$, $1 \le p < \infty$.

on $[0,1)$ and 0 at 1. When $n \ge 2$, define h_n by letting m be the positive integer such that $2^{m-1} < n \le 2^m$, then let

$$h_n(t) = \begin{cases} 1 & \text{if } \frac{2n-2}{2^m} - 1 \le t < \frac{2n-1}{2^m} - 1; \\ -1 & \text{if } \frac{2n-1}{2^m} - 1 \le t < \frac{2n}{2^m} - 1; \\ 0 & \text{otherwise.} \end{cases}$$

See Figure 4.2 for the graphs of the first few of these functions, which will make it clear how they are being constructed. The vertical dashed lines in the graphs are included for clarity. Notice that each h_n is a positive multiple of the derivative of the corresponding member s_n of the classical Schauder basis for $C[0,1]$.

Suppose that (α_n) is a sequence of scalars, that n_0 and m_0 are positive integers such that $n_0 \ge 2$ and $2^{m_0-1} < n_0 \le 2^{m_0}$, and that I_1 and I_2 are the respective intervals $[\frac{2n_0-2}{2^{m_0}} - 1, \frac{2n_0-1}{2^{m_0}} - 1)$ and $[\frac{2n_0-1}{2^{m_0}} - 1, \frac{2n_0}{2^{m_0}} - 1)$. Let α be the constant value of $\sum_{n=1}^{n_0-1} \alpha_n h_n$ on $I_1 \cup I_2$. It is a straightforward calculus exercise to show that $s^p + t^p - 2\big((s+t)/2\big)^p \ge 0$ when $s,t \ge 0$, from which it follows that

$$\begin{aligned} \int_{[0,1]} \Big|\sum_{n=1}^{n_0} \alpha_n h_n\Big|^p - \int_{[0,1]} \Big|\sum_{n=1}^{n_0-1} \alpha_n h_n\Big|^p &= \int_{I_1} |\alpha + \alpha_{n_0}|^p + \int_{I_2} |\alpha - \alpha_{n_0}|^p - \int_{I_1 \cup I_2} |\alpha|^p \\ &= \frac{|\alpha + \alpha_{n_0}|^p + |\alpha - \alpha_{n_0}|^p - 2|\alpha|^p}{2^{m_0}} \\ &\ge 0. \end{aligned}$$

Therefore $\|\sum_{n=1}^{n_0-1} \alpha_n h_n\|_p \le \|\sum_{n=1}^{n_0} \alpha_n h_n\|_p$, and so $\|\sum_{n=1}^{m_1} \alpha_n h_n\|_p \le \|\sum_{n=1}^{m_2} \alpha_n h_n\|_p$ whenever $m_1, m_2 \in \mathbb{N}$ and $m_1 \le m_2$.

It is easy to check that $\langle\{\, h_n : n \in \mathbb{N}\,\}\rangle$ contains the indicator function of every dyadic interval $[\frac{n-1}{2^m}, \frac{n}{2^m})$ such that m is a nonnegative integer and n is an integer for which $1 \le n \le 2^m$. It follows that $L_p[0,1] = [\{ h_n : n \in \mathbb{N} \}]$; see the argument of Example 1.12.3. Since each h_n is nonzero, an application of Theorem 4.1.24 shows that (h_n) is a basis for $L_p[0,1]$. Notice that this basis is monotone by the equivalence of (a) and (c) in Proposition 4.1.21.

The sequence (h_n) is called the *Haar basis* for $L_p[0,1]$. The fact that it is a basis for $L_p[0,1]$ was first shown by Schauder in a 1928 paper [207].

The infinite-dimensional separable Banach spaces are, up to isometric isomorphism, just the infinite-dimensional closed subspaces of $C[0,1]$; see the comments preceding Example 4.1.11. Since $C[0,1]$ has a basis, it is natural to ask if every infinite-dimensional separable Banach space has a basis. This question appeared in Banach's book [13, p. 111] and is called Banach's *basis problem.* It remained open for forty years, but was finally settled in the negative in a 1973 paper by Per Enflo [76], who found a reflexive counterexample. The space is the same one used by Enflo as an example of a Banach space lacking the approximation property; see the discussion following Theorem 3.4.32. The connection between bases and the approximation property will be explored briefly at the end of this section.

It is at least true that every infinite-dimensional Banach space has a basic sequence in it. This result has a rather interesting history. It is often attributed to Banach since it first appeared in his book [13, p. 238], without proof, as an afterthought to some remarks on his basis problem. The way it is stated there almost makes it sound as if Banach is leaving the result as a straightforward exercise for the reader.[1] However, no published proof appeared for twenty-six years. Proving this result was an idea whose time had apparently come in 1958, for three proofs appeared that year, by Bernard Gelbaum [86], Czesław Bessaga and Aleksander Pełczyński [26], and Mahlon Day [53, p. 72] (though Day's proof contained an error that he corrected in a 1962 paper [55]). It is a testament to the faith put in Banach that the review of [26] in *Mathematical Reviews* referred to Bessaga and Pełczyński's accomplishment as a new proof of a previously known result.

Though this leaves open the question of whether Banach really did have a proof, a 1962 paper by Pełczyński [179] provides a clue. In that paper, Pełczyński displayed a method of constructing basic sequences that he attributed to S. Mazur and that can be used to produce a fairly simple proof that infinite-dimensional Banach spaces always have basic sequences in them. It seems likely that Banach knew of Mazur's method and had a proof based on that method in mind when he made his claim.

The essence of Mazur's method is contained in the next two lemmas. Suppose that $n \in \mathbb{N}$, that Y is a proper subspace of Euclidean n-space $\mathbb{F}^n$, and that x_0 is a member of $\mathbb{F}^n$ such that $\|y\| \le \|y + \alpha x_0\|$ whenever

---

[1] "Remarquons toutefois que tout espace du type (B) à une infinité de dimensions renferme un ensemble linéaire fermé à une infinité de dimensions qui admet une base."

$y \in Y$ and $\alpha \in \mathbb{F}$. Then x_0 is orthogonal to Y. The second lemma says that, analogously, for every finite-dimensional subspace Y of an infinite-dimensional Banach space X there is a norm-one vector in X that is "almost orthogonal" to Y. To prove that every infinite-dimensional Banach space has a basic sequence in it, the general plan of attack is to use this lemma to construct a particular sequence (x_n) of norm-one members of the Banach space such that x_n is, in the sense of the lemma, "almost orthogonal" to $\langle x_1, \ldots, x_{n-1} \rangle$ when $n > 1$, then show that this sequence is basic.

The only immediate use made of the first lemma is to derive the second, and in fact the second lemma could be proved directly by a slightly modified version of the proof of the first. However, the first will have a further application later in this section, in the proof of Theorem 4.1.32.

4.1.28 Lemma. *Suppose that Y is a nontrivial finite-dimensional subspace of a Banach space X and that $0 < \epsilon < 1$. Let $y_1, \ldots, y_m$ be elements of the compact set S_Y such that every member of S_Y is within $\epsilon/4$ of some y_j, and let $x_1^*, \ldots, x_m^*$ be members of S_{X^*} such that $x_j^* y_j = 1$ for each j. Suppose that (x_n) is a sequence in X such that $\inf_n \|x_n\| > 0$ and $\lim_n x_j^* x_n = 0$ when $j = 1, \ldots, m$. Then for every positive integer N there is a positive integer $n_{\epsilon,N}$ greater than N such that $\|y\| \le (1+\epsilon)\|y + \alpha x_{n_{\epsilon,N}}\|$ whenever $y \in Y$ and $\alpha \in \mathbb{F}$.*

PROOF. Suppose that $N \in \mathbb{N}$. Let $n_{\epsilon,N}$ be a positive integer greater than N such that $|x_j^* x_{n_{\epsilon,N}}| < (\epsilon \inf_n \|x_n\|)/8$ when $j = 1, \ldots, m$. Since $1 - \epsilon/2 > 1/(1+\epsilon)$, it is enough to prove that $\|y + \alpha x_{n_{\epsilon,N}}\| > 1 - \epsilon/2$ whenever $y \in S_Y$ and $\alpha \in \mathbb{F}$.

Suppose that $y \in S_Y$ and $\alpha \in \mathbb{F}$. If $|\alpha| \ge 2/\|x_{n_{\epsilon,N}}\|$, then $\|y + \alpha x_{n_{\epsilon,N}}\| \ge \|\alpha x_{n_{\epsilon,N}}\| - \|y\| \ge 1 > 1 - \epsilon/2$, so it may be assumed that $|\alpha| < 2/\|x_{n_{\epsilon,N}}\|$. Since there is a j_0 such that $\|y - y_{j_0}\| < \epsilon/4$, it follows that

$$\begin{aligned}
\|y + \alpha x_{n_{\epsilon,N}}\| &> \|y_{j_0} + \alpha x_{n_{\epsilon,N}}\| - \frac{\epsilon}{4} \\
&\ge |x_{j_0}^*(y_{j_0} + \alpha x_{n_{\epsilon,N}})| - \frac{\epsilon}{4} \\
&\ge |x_{j_0}^* y_{j_0}| - |\alpha|\,|x_{j_0}^* x_{n_{\epsilon,N}}| - \frac{\epsilon}{4} \\
&> 1 - \frac{\epsilon}{4} - \frac{\epsilon}{4} \\
&= 1 - \frac{\epsilon}{2},
\end{aligned}$$

as required. ■

4.1.29 Lemma. *Suppose that Y is a finite-dimensional subspace of an infinite-dimensional Banach space X. Then for every positive ϵ there is an x_ϵ in S_X such that $\|y\| \le (1+\epsilon)\|y + \alpha x_\epsilon\|$ whenever $y \in Y$ and $\alpha \in \mathbb{F}$.*

PROOF. Fix a positive ϵ. It may be assumed that $\epsilon < 1$ and that $Y \ne \{0\}$. Let $y_1, \ldots, y_m$ and $x_1^*, \ldots, x_m^*$ be as in the statement of Lemma 4.1.28.

Since not every member of X^* is a linear combination of $x_1^*, \ldots, x_m^*$, it follows from Lemma 1.9.11 that $\bigcap_{j=1}^m \ker x_j^* \neq \{0\}$. Let x_ϵ be any member of $\bigcap_{j=1}^m \ker x_j^*$ having norm 1 and let $x_n = x_\epsilon$ for each n. Then Lemma 4.1.28 assures that $\|y\| \leq (1+\epsilon)\|y + \alpha x_\epsilon\|$ whenever $y \in Y$ and $\alpha \in \mathbb{F}$. ■

4.1.30 Theorem. *Suppose that X is an infinite-dimensional Banach space and that $M > 1$. Then there is a normalized basic sequence in X having basis constant no more than M.*

PROOF. Let x_1 be any member of S_X. By Lemma 4.1.29, there an x_2 in S_X such that $\|\alpha_1 x_1\| \leq M^{1/2}\|\alpha_1 x_1 + \alpha_2 x_2\|$ whenever $\alpha_1, \alpha_2 \in \mathbb{F}$. Further applications of the lemma then produce a sequence (x_n) in S_X such that $\|\sum_{n=1}^m \alpha_n x_n\| \leq M^{1/2^m}\|\sum_{n=1}^{m+1} \alpha_n x_n\|$ whenever m is a positive integer and $\alpha_1, \ldots, \alpha_{m+1}$ are scalars. It follows that

$$\left\|\sum_{n=1}^{m_1} \alpha_n x_n\right\| \leq \left(\prod_{n=m_1}^{m_2-1} M^{1/2^n}\right)\left\|\sum_{n=1}^{m_2} \alpha_n x_n\right\| \leq M\left\|\sum_{n=1}^{m_2} \alpha_n x_n\right\|$$

whenever $m_1, m_2 \in \mathbb{N}$, $m_1 < m_2$, and $\alpha_1, \ldots, \alpha_{m_2} \in \mathbb{F}$. The sequence (x_n) is therefore basic by Corollary 4.1.25 and has basis constant no more than M by Proposition 4.1.20. ■

Thus, every infinite-dimensional Banach space has in it basic sequences with basis constants as close to 1 as desired. This naturally leads to the question of whether every Banach space with a basis must have bases with basis constants arbitrarily close to 1; that is, whether the following quantity must be 1 for every Banach space possessing a basis.

4.1.31 Definition. Suppose that X is a Banach space having a basis. For each basis (x_n) of X, let $K_{(x_n)}$ be the basis constant for (x_n). Then the *basis constant* for X is $\inf\{\, K_{(x_n)} : (x_n) \text{ is a basis for } X\,\}$.

Of course, if every Banach space with a basis were to have a monotone basis, then every Banach space with a basis would have basis constant 1. However, it was shown in a series of papers by V. I. Gurariĭ, culminating in the 1965 paper [97], that some Banach spaces with bases do not have monotone bases; see pages 241–248 and 623 of [216] for an exposition of Gurariĭ's result. The larger problem was put to rest by Per Enflo in a 1973 paper [75] in which he constructed a Banach space having a basis such that the basis constant of the space is greater than 1.

Though Theorem 4.1.30 guarantees the presence of basic sequences inside every infinite-dimensional Banach space, it does not suggest any specific places in a Banach space to look for them. The following result does. It is essentially this result that Pełczyński obtained in [179] using Mazur's method;

see the comments preceding Lemma 4.1.28. In his paper, Pełczyński points out that the result was first proved by Czesław Bessaga in his thesis, also using Mazur's method.

Notice the close similarity between the proof of this theorem and that of Theorem 4.1.30.

4.1.32 Theorem. *Suppose that (x_n) is a sequence in a Banach space such that (x_n) converges weakly to zero but does not converge to zero with respect to the norm topology. Then some subsequence of (x_n) is a basic sequence.*

PROOF. Since some subsequence of (x_n) is bounded away from 0, it may be assumed that $\inf_n \|x_n\| > 0$. Let M be any real number greater than 1, and let $n_1 = 1$. By Lemma 4.1.28, there is a positive integer n_2 greater than n_1 such that $\|\alpha_1 x_{n_1}\| \le M^{1/2}\|\alpha_1 x_{n_1} + \alpha_2 x_{n_2}\|$ whenever $\alpha_1, \alpha_2 \in \mathbb{F}$. An easy induction argument based on Lemma 4.1.28 then produces a subsequence (x_{n_j}) of (x_n) such that $\|\sum_{j=1}^m \alpha_j x_{n_j}\| \le M^{1/2^m}\|\sum_{j=1}^{m+1} \alpha_j x_{n_j}\|$ whenever $m \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_{m+1} \in \mathbb{F}$. It follows that

$$\left\|\sum_{j=1}^{m_1} \alpha_j x_{n_j}\right\| \le \left(\prod_{j=m_1}^{m_2-1} M^{1/2^j}\right)\left\|\sum_{j=1}^{m_2} \alpha_j x_{n_j}\right\| \le M\left\|\sum_{j=1}^{m_2} \alpha_j x_{n_j}\right\|$$

whenever $m_1, m_2 \in \mathbb{N}$, $m_1 < m_2$, and $\alpha_1, \ldots, \alpha_{m_2} \in \mathbb{F}$. The sequence (x_{n_j}) is therefore basic by Corollary 4.1.25. ■

The final topic for this section is the brief exploration of the connection between bases and the approximation property promised earlier. The main result in this direction is the following one.

4.1.33 Theorem. *Every Banach space with a basis has the approximation property.*

PROOF. Suppose that X is a Banach space with a basis (x_n), that (P_n) is the sequence of natural projections for (x_n), and that M is the basis constant for (x_n). Let K be a nonempty compact subset of X and let ϵ be a positive number. By Theorem 3.4.32, it is enough to find a positive integer n_0 such that $\|P_{n_0}x - x\| < \epsilon$ whenever $x \in X$.

Let $y_1, \ldots, y_m$ be members of K such that every member of K is within distance $\epsilon/(M+2)$ of some y_j, and let n_0 be a positive integer such that $\|P_{n_0}y_j - y_j\| < \epsilon/(M+2)$ when $j = 1, \ldots, m$. Suppose that x is any member of K. Let j_0 be such that $\|x - y_{j_0}\| < \epsilon/(M+2)$. Then

$$\|P_{n_0}x - x\| \le \|P_{n_0}(x - y_{j_0})\| + \|P_{n_0}y_{j_0} - y_{j_0}\| + \|y_{j_0} - x\| < \epsilon,$$

which finishes the proof. ■

Thus, Enflo's example of an infinite-dimensional separable Banach space lacking the approximation property is also an example of an infinite-dimensional separable Banach space lacking a basis.

Actually, a bit more was proved in the preceding theorem than advertised. In the notation of the theorem, it was shown not only that the identity operator on X can be uniformly approximated on compact subsets of X by bounded finite-rank linear operators, but also that the family of approximating operators can be selected to be a bounded subset of $B(X)$, since $\sup_n \|P_n\|$ is the finite basis constant for (x_n). That is, it was shown that every Banach space with a basis has the following property.

4.1.34 Definition. Let X be a Banach space. Suppose that there is a positive constant t having the property that, for every compact subset K of X and every positive ϵ, there is a finite-rank member $T_{K,\epsilon}$ of $B(X)$ such that $\|T_{K,\epsilon}\| \leq t$ and $\|T_{K,\epsilon}x - x\| < \epsilon$ whenever $x \in K$. Then X has the *bounded approximation property.*

A Banach space that satisfies the conditions of the preceding definition for a specific t is said to have the *t-approximation property,* and is said to have the *metric approximation property* if it has the 1-approximation property. See Exercise 4.21 for an important characterization of the bounded approximation property for separable Banach spaces.

Though the bounded approximation property obviously implies the approximation property, it was shown by Tadeusz Figiel and William B. Johnson in a 1973 paper [78] that there are infinite-dimensional separable Banach spaces with the approximation property but not the bounded approximation property. It is therefore of some interest that every Banach space with a basis has not just the approximation property, but in fact the bounded approximation property.

Incidentally, there are infinite-dimensional separable Banach spaces having the bounded approximation property but no basis. The first example of one appeared in a 1987 paper of Stanisław Szarek [230]. Thus, for infinite-dimensional separable Banach spaces the property of having a basis implies the bounded approximation property, which in turn implies the approximation property, but neither of the implications is reversible.

For a further discussion of the relationship between the approximation property, the bounded approximation property, and the existence of bases, see the comments of Pełczyński and Bessaga in either [14] or [15].

Exercises

4.1 Suppose that $\mathfrak{B}$ is a vector space basis for an infinite-dimensional Banach space and that $\mathfrak{B}^{\#}$ is the collection of coordinate functionals for $\mathfrak{B}$. The purpose of this exercise is to show that not all members of $\mathfrak{B}^{\#}$ can be bounded, and that, in fact, only finitely many can be so. Let $\mathfrak{B}_b^\#$ be the collection of members of $\mathfrak{B}^\#$ that are bounded.

(a) Show by example that $\mathfrak{B}_b^\#$ might be nonempty.

(b) Show that $\sup\{ \|f\| : f \in \mathfrak{B}_b^\# \} \neq +\infty$.

(c) Show that $\mathfrak{B}_b^\#$ is finite.

4.2 Give an example of an infinite-dimensional normed space with a vector space basis $\mathfrak{B}$ such that every coordinate functional for $\mathfrak{B}$ is bounded. (Compare Exercise 4.1.)

4.3 Prove the claims made in Example 4.1.3.

4.4 Prove that the sequence $(1, 1, 1, \dots), e_1, e_2, e_3, \dots$ is a basis for the Banach space c of Exercise 1.25.

4.5 Prove that if Banach spaces X and Y have bases, then so does $X \oplus Y$.

4.6 Suppose that (x_n) is a basis for a Banach space X and that (n_j) is a sequence of positive integers such that $\mathbb{N} \setminus \{ n_j : j \in \mathbb{N} \}$ is infinite. Let $M = [\{ x_{n_j} : j \in \mathbb{N} \}]$. Prove that both M and X/M have bases.

4.7 Suppose that X is a complex Banach space and X_r is the real Banach space obtained from X by restricting multiplication of vectors by scalars to $\mathbb{R} \times X$. Show that a sequence (x_n) in X is a basis for X if and only if the sequence $(x_1, ix_1, x_2, ix_2, \dots)$ is a basis for X_r.

4.8 Suppose that (x_n) is a basis for a Banach space and that (λ_n) is a sequence of nonzero scalars.

(a) Must each coordinate functional for $(\lambda_n x_n)$ be the same as the corresponding coordinate functional for (x_n)?

(b) Must each natural projection for $(\lambda_n x_n)$ be the same as the corresponding natural projection for (x_n)?

(c) Must the basis constant for $(\lambda_n x_n)$ be the same as the basis constant for (x_n)?

4.9 Prove that the classical Schauder basis for $C[0,1]$ is monotone. Is it strictly monotone?

4.10 Suppose that $1 \leq p < \infty$. Is the Haar basis for $L_p[0,1]$ strictly monotone for any or all such p?

4.11 Prove that every Banach space with a basis has a nonmonotone basis.

4.12 (a) Suppose that (x_n) is a basis for a Banach space X. What is wrong with the following "proof" that there is only one norm for X equivalent to the original norm of X with respect to which (x_n) is monotone? "By the equivalence of (a) and (d) in Proposition 4.1.21, every norm for X equivalent to the original norm of X with respect to which (x_n) is monotone is the same as the (x_n) norm of X, so there is only one such norm."

(b) Display a basis (x_n) for a Banach space X and two different norms for X equivalent to the original norm of X with respect to which (x_n) is monotone.

4.13 The purpose of this exercise is to show that uniqueness of expansions is not enough to guarantee that a sequence in a Banach space is basic; that is, that it is possible for a sequence (x_n) in a Banach space not to be basic even though the sequences (α_n) and (β_n) of scalars are the same whenever $\sum_n \alpha_n x_n$ and $\sum_n \beta_n x_n$ both converge and are equal. Show this by considering the sequence $(x_n)_{n=0}^{\infty}$ in $C[0,1]$ defined by the formulas $x_0(t) = 1$ and $x_n(t) = t^n$ when $n \geq 1$.

4.14 Suppose that (x_n) is a basis for a Banach space and that (x_n^*) is the sequence of coordinate functionals for (x_n). Prove that there is a positive constant M such that $1 \leq \|x_n\|\,\|x_n^*\| \leq M$ for each n.

4.15 Suppose that (x_n) is a basis for a Banach space and that (x_n^*) is the sequence of coordinate functionals for (x_n).

(a) Show by example that $\sup_n \|x_n^*\|$ might not be finite.

(b) Find necessary and sufficient conditions on (x_n) for $\sup_n \|x_n^*\|$ to be finite. (Exercise 4.14 might help.)

(c) Find necessary and sufficient conditions on (x_n^*) for $\sup_n \|x_n\|$ to be finite.

4.16 Suppose that (x_n) is a basis for a Banach space.

(a) Show by example that it might not be true that $\sum_n \alpha_n x_n$ converges whenever $\sum_n |\alpha_n| < \infty$.

(b) Find necessary and sufficient conditions on (x_n) for $\sum_n \alpha_n x_n$ to be convergent whenever $\sum_n |\alpha_n| < \infty$.

4.17 Suppose that (x_n) is a basic sequence in a Banach space.

(a) Prove that 0 is the only possible weak limit point of $\{x_n : n \in \mathbb{N}\}$.

(b) Show by example that $\{x_n : n \in \mathbb{N}\}$ might have 0 as a weak limit point, even if (x_n) is normalized.

(c) Show by example that $\{x_n : n \in \mathbb{N}\}$ might have no weak limit points whatever, even if (x_n) is normalized.

4.18 Suppose that $x_1, \ldots, x_m$ is a nonempty linearly independent finite list in an infinite-dimensional Banach space. Show that this list can be extended to a basic sequence (x_n).

4.19 Theorem 4.1.32 is not very useful for finding basic sequences inside ℓ_1. Why not?

4.20 Give a proof of Theorem 4.1.33 that uses nothing deeper about the approximation property than its definition, and in particular avoids the use of Theorem 3.4.32.

4.21 Prove that a separable Banach space X has the bounded approximation property if and only if there is a sequence (T_n) of finite-rank members of $B(X)$ such that $\lim_n T_n x = x$ for each x in X.

4.22 Prove that every Banach space with a monotone basis has the metric approximation property.

4.23 The major content of Theorem 4.1.24—namely, that conditions (1), (2), and (3) in the statement of that theorem together imply that the sequence (x_n) is a basis—does appear in Banach's book [13] in an equivalent form, but it is not all that easy to spot. Banach's result is actually about *biorthogonal sequences* rather than bases, where a biorthogonal "sequence" is actually a pair of sequences $(x_n), (x_n^*)$ such that (x_n) lies in some Banach space, the sequence (x_n^*) lies in the dual of that space, and $x_m^* x_n$ is 1 if $m = n$ and 0 otherwise. This is Banach's result, which appears as Theorem 4 of Chapter VII on page 108 of his book:

Suppose that X is a Banach space; that (x_n) is a sequence in X and (x_n^*) a sequence in X^* that together form a biorthogonal sequence; that $[\{ x_n : n \in \mathbb{N} \}] = X$; and that, for each x in X, the partial sums of the formal series $\sum_n (x_n^* x) x_n$ are bounded. Then $\sum_n (x_n^* x) x_n$ converges for each x in X.

Here is another way to look at this result. Suppose that X, (x_n), and (x_n^*) satisfy the initial parts of the hypotheses of Banach's result up to and including the requirement that $[\{ x_n : n \in \mathbb{N} \}] = X$. For each positive integer m, let P_m be the member of $B(X)$ given by the formula $P_m x = \sum_{n=1}^{m} (x_n^* x) x_n$. Banach's result says that if $\sup_m \|P_m x\|$ is finite for each x in X, then $\lim_m P_m x$ exists for each x in X. The purpose of this exercise is to demonstrate that Banach's result is essentially the same as the major content of Theorem 4.1.24 by showing that each can be readily derived from the other.

(a) Suppose that X, (x_n), and (x_n^*) satisfy the hypotheses of Banach's result. Use Theorem 4.1.24 to show that the sequence (x_n) is a basis for X having (x_n^*) as its sequence of coordinate functionals, and that the conclusion of Banach's result follows.

For the rest of this exercise, assume that Banach's result is known to hold but that neither Theorem 4.1.24 nor the consequences of it demonstrated in part (a) have been proved.

(b) Show that under the hypotheses of Banach's result and using the notation in the explanation of Banach's result given above, it follows that $\lim_m P_m x = x$ whenever $x \in X$, that is, that $x = \sum_n (x_n^* x) x_n$ for each x in X.

(c) Suppose that (x_n) is a sequence in a Banach space X satisfying conditions (1), (2), and (3) of Theorem 4.1.24. Use Banach's result to prove that (x_n) is a basis for X. You may duplicate the relevant portion of the proof of Theorem 4.1.24 up to and including the construction of the maps P_m, but you should then concentrate on obtaining the sequence (x_n^*) as in the hypotheses of Banach's result with the goal of applying his result as quickly thereafter as possible.

4.2 Unconditional Bases

The following two propositions will prove useful in what is to follow. Though these results were not needed until now, their proofs are entirely elementary and could have been given in Section 1.3. The first of these propositions has a much shorter, though less elementary, proof; see Exercise 4.25.

4.2.1 Proposition. *If $\sum_n x_n$ is an unconditionally convergent series in a normed space, then $\sum_n x_{\pi(n)} = \sum_n x_n$ for each permutation π of $\mathbb{N}$.*

PROOF. Suppose that $\sum_n x_n$ is a series in a normed space and that π is a permutation of $\mathbb{N}$ such that $\sum_n x_n$ and $\sum_n x_{\pi(n)}$ both converge but to different limits. It will be shown that there is another permutation π' of $\mathbb{N}$ such that $\sum_n x_{\pi'(n)}$ does not converge, from which it follows that $\sum_n x_n$ is only conditionally convergent.

Let $\epsilon = \|\sum_n x_{\pi(n)} - \sum_n x_n\|$ and let the positive integer p_1 be such that

$$\left\|\sum_n x_{\pi(n)} - \sum_{n=1}^{p_1} x_{\pi(n)}\right\| < \frac{\epsilon}{3}.$$

There is a positive integer q_1 such that

$$\{\,\pi(n) : n \in \mathbb{N},\ 1 \le n \le p_1\,\} \subseteq \{\,n : n \in \mathbb{N},\ 1 \le n \le q_1\,\}$$

and

$$\left\|\sum_n x_n - \sum_{n=1}^{q_1} x_n\right\| < \frac{\epsilon}{3}.$$

There is then a positive integer p_2 such that

$$\{\,n : n \in \mathbb{N},\ 1 \le n \le q_1\,\} \subseteq \{\,\pi(n) : n \in \mathbb{N},\ 1 \le n \le p_2\,\}$$

and

$$\left\|\sum_n x_{\pi(n)} - \sum_{n=1}^{p_2} x_{\pi(n)}\right\| < \frac{\epsilon}{3},$$

and a further positive integer q_2 such that

$$\{\,\pi(n) : n \in \mathbb{N},\ 1 \le n \le p_2\,\} \subseteq \{\,n : n \in \mathbb{N},\ 1 \le n \le q_2\,\}$$

and

$$\left\|\sum_n x_n - \sum_{n=1}^{q_2} x_n\right\| < \frac{\epsilon}{3}.$$

Continue the construction of the sequences (p_n) and (q_n) in this fashion.

Now let π' be the permutation of $\mathbb{N}$ obtained by listing $\mathbb{N}$ in the following order. First list $\pi(1)$ through $\pi(p_1)$, then follow this by the members of $1, \ldots, q_1$ not already listed. Follow this by the members of $\pi(1), \ldots, \pi(p_2)$

not already listed, and in turn follow that by the members of $1, \ldots, q_2$ not already listed, and so forth. Since the partial sums of $\sum_n x_{\pi'(n)}$ swing back and forth between being within $\epsilon/3$ of $\sum_n x_{\pi(n)}$ and being within $\epsilon/3$ of $\sum_n x_n$, the series $\sum_n x_{\pi'(n)}$ does not converge. ∎

4.2.2 Definition. In any setting in which the notion of series makes sense, a *subseries* of a formal series $\sum_n x_n$ is a formal series $\sum_j x_{n_j}$ obtained from a subsequence (x_{n_j}) of (x_n).

4.2.3 Proposition. *A formal series $\sum_n x_n$ in a Banach space is unconditionally convergent if and only if each subseries of $\sum_n x_n$ converges.*

PROOF. Suppose first that there is a subsequence (x_{n_j}) of (x_n) such that $\sum_j x_{n_j}$ does not converge. Then there must be a positive ϵ and sequences (p_n) and (q_n) of positive integers such that

$$p_1 \le q_1 < p_2 \le q_2 < p_3 \le q_3 < \cdots$$

and $\|\sum_{j=p_k}^{q_k} x_{n_j}\| \ge \epsilon$ for each k. It may be assumed that infinitely many positive integers are omitted from $\bigcup_k \{ n_j : p_k \le j \le q_k \}$. Let (r_n) be the sequence consisting of those omitted integers in ascending order, and let π be the permutation of $\mathbb{N}$ into the order

$$n_{p_1}, \ldots, n_{q_1}, r_1, n_{p_2}, \ldots, n_{q_2}, r_2, n_{p_3}, \ldots, n_{q_3}, r_3, \ldots$$

where the admittedly ambiguous notation $n_{p_k}, \ldots, n_{q_k}$ is being used here to represent the particular terms of (n_j) that are indexed by p_k through q_k rather than all integers j such that $n_{p_k} \le j \le n_{q_k}$. It is clear that $\sum_n x_{\pi(n)}$ does not converge, so $\sum_n x_n$ is not unconditionally convergent.

Suppose conversely that $\sum_n x_{\pi(n)}$ does not converge for some permutation π of $\mathbb{N}$. Then there must be a positive ϵ and sequences (s_n) and (t_n) of positive integers such that

$$s_1 \le t_1 < s_2 \le t_2 < s_3 \le t_3 < \cdots$$

and $\|\sum_{n=s_k}^{t_k} x_{\pi(n)}\| \ge \epsilon$ for each k. It may be assumed that

$$\max\{ \pi(n) : s_k \le n \le t_k \} < \min\{ \pi(n) : s_{k+1} \le n \le t_{k+1} \}$$

for each k. Let (n_j) be formed by applying π to the terms of the sequence $s_1, \ldots, t_1, s_2, \ldots, t_2, \ldots$ to get $\pi(s_1), \ldots, \pi(t_1), \pi(s_2), \ldots, \pi(t_2), \ldots$ and sorting the resulting sequence into ascending order. Then the subseries $\sum_j x_{n_j}$ of $\sum_n x_n$ does not converge. ∎

4.2.4 Corollary. *If a series in a Banach space converges unconditionally, then so does each of its subseries.*

An important consequence of Proposition 4.2.1 and Corollary 4.2.4 is that it is possible to define "unordered subseries" of unconditionally convergent series in Banach spaces in the following way.

4.2.5 Definition. Suppose that $\sum_n x_n$ is an unconditionally convergent series in a Banach space X and that $A \subseteq \mathbb{N}$. If $A = \varnothing$, then $\sum_{n\in A} x_n$ is the zero element of X, otherwise $\sum_{n\in A} x_n$ is the member of X obtained by listing the elements of A in any order $n_1, n_2, \ldots$ and letting $\sum_{n\in A} x_n = x_{n_1} + x_{n_2} + \cdots$.

One other important characterization of unconditional convergence in Banach spaces requires a bit of preliminary work.

4.2.6 Lemma. *If Y is a closed subspace of ℓ_∞ that contains every sequence of scalars whose terms all come from $\{0, 1\}$, then $Y = \ell_\infty$.*

PROOF. Suppose that the member (t_n) of S_{ℓ_∞} has only nonnegative real terms. It is enough to show that $(t_n) \in Y$. For each positive integer n, let $0.s_{n,1}s_{n,2}s_{n,3}\cdots$ be a binary expansion of t_n. Then $\big((s_{n,j})\big)_{j=1}^\infty$ is a sequence in Y, and $(t_n) = \sum_j 2^{-j}(s_{n,j}) \in Y$. ■

The next lemma is proved in a bit more generality than is really needed here, but the more general version will have a further use in Section 4.3 to prove Proposition 4.3.9.

4.2.7 Lemma. *Suppose that $\sum_n x_n$ is a formal series in a normed space X such that $\sum_n x^*x_n$ is absolutely convergent whenever $x^* \in X^*$. Then there is a nonnegative real number M such that $\sup_m \|\sum_{n=1}^m \alpha_n x_n\| \le M\|(\alpha_n)\|_\infty$ whenever $(\alpha_n) \in \ell_\infty$.*

PROOF. Define $T\colon X^* \to \ell_1$ by the formula $T(x^*) = (x^*x_n)$. It is clear that T is linear, and it follows from the closed graph theorem that T is bounded. Fix a member (α_n) of ℓ_∞. If $m \in \mathbb{N}$ and $x^* \in B_{X^*}$, then

$$\left|x^*\left(\sum_{n=1}^m \alpha_n x_n\right)\right| = \left|\sum_{n=1}^m \alpha_n x^* x_n\right| \le \|(\alpha_n)\|_\infty \|Tx^*\|_1 \le \|(\alpha_n)\|_\infty \|T\|,$$

so $\|\sum_{n=1}^m \alpha_n x_n\| \le \|T\|\,\|(\alpha_n)\|_\infty$ whenever $m \in \mathbb{N}$. Let $M = \|T\|$ to finish the proof. ■

4.2.8 The Bounded Multiplier Test. *A formal series $\sum_n x_n$ in a Banach space is unconditionally convergent if and only if $\sum_n \alpha_n x_n$ converges whenever $(\alpha_n) \in \ell_\infty$.*

PROOF. Let X be the Banach space in which the sequence (x_n) lies. Suppose first that $\sum_n \alpha_n x_n$ converges whenever $(\alpha_n) \in \ell_\infty$. It follows immediately that every subseries of $\sum_n x_n$ converges, so $\sum_n x_n$ is unconditionally convergent.

Suppose conversely that $\sum_n x_n$ is unconditionally convergent. Let Y be the subspace of ℓ_∞ consisting of all members (α_n) of ℓ_∞ such that $\sum_n \alpha_n x_n$ converges. The goal is to show that $Y = \ell_\infty$. Since each subseries of $\sum_n x_n$ converges, every sequence of scalars whose terms all come from $\{0, 1\}$ is in Y, so by Lemma 4.2.6 it is enough to show that Y is closed.

Suppose that $\big((\alpha_{n,j})\big)_{j=1}^\infty$ is a sequence in Y that converges to some (α_n) in ℓ_∞. It is enough to show that $(\alpha_n) \in Y$, for which it is enough to show that the sequence of partial sums of $\sum_n \alpha_n x_n$ is Cauchy. To this end, suppose that $\epsilon > 0$. For each x^* in X^*, the series $\sum_n x^* x_n$ is a series of scalars that is, by Proposition 4.2.3, unconditionally convergent and therefore absolutely convergent, so $\sum_n x_n$ satisfies the hypotheses of Lemma 4.2.7. Let M be as in the conclusion of that lemma. Let j_ϵ be a positive integer such that $\|(\alpha_{n,j_\epsilon}) - (\alpha_n)\|_\infty < \epsilon$, then let N be a positive integer such that $\|\sum_{n=p}^q \alpha_{n,j_\epsilon} x_n\| < \epsilon$ whenever $q \geq p \geq N$. It follows that if $q \geq p \geq N$, then

$$\begin{aligned}\left\|\sum_{n=p}^q \alpha_n x_n\right\| &\leq \left\|\sum_{n=p}^q \alpha_{n,j_\epsilon} x_n\right\| + \left\|\sum_{n=p}^q (\alpha_{n,j_\epsilon} - \alpha_n) x_n\right\| \\ &< \epsilon + M\|(\alpha_{n,j_\epsilon}) - (\alpha_n)\|_\infty \\ &\leq (1+M)\epsilon.\end{aligned}$$

Thus, the sequence of partial sums of $\sum_n \alpha_n x_n$ is Cauchy. ■

4.2.9 Corollary. *If a series $\sum_n x_n$ in a Banach space converges unconditionally, then so does $\sum_n \alpha_n x_n$ whenever $(\alpha_n) \in \ell_\infty$.*

Considering the title of this section and the emphasis already placed on unconditionally convergent series, the reader may have anticipated the following definition.

4.2.10 Definition. A basis (x_n) for a Banach space X is *unconditional* if, for every x in X, the expansion $\sum_n \alpha_n x_n$ for x in terms of the basis is unconditionally convergent. A basis for a Banach space is *conditional* if it is not unconditional.

The term *unconditional basis* is due to R. C. James [109]. However, others studied such bases before James gave them their modern name in his 1950 paper. In particular, S. Karlin [130] called such bases *absolute* in a study published two years earlier, and this term was still used by some authors at least as late as 1962; see, for example, [82].

4.2.11 Example. Suppose that X is c_0 or ℓ_p such that $1 \leq p < \infty$. Then the standard unit vector basis for X is clearly unconditional.

The following elementary results about unconditional bases follow easily from the corresponding general results about bases in Section 4.1.

4.2.12 Proposition. *Suppose that* (x_n) *is an unconditional basis for a Banach space* X *and that* (λ_n) *is a sequence of nonzero scalars. Then* $(\lambda_n x_n)$ *is also an unconditional basis for* X.

4.2.13 Corollary. *If* (x_n) *is an unconditional basis for a Banach space, then* $(\|x_n\|^{-1}x_n)$ *is a normalized unconditional basis for the same space.*

4.2.14 Proposition. *Suppose that* X *and* Y *are Banach spaces, that* T *is an isomorphism from* X *into* Y, *and that* (x_n) *is an unconditional basic sequence in* X. *Then* (Tx_n) *is an unconditional basic sequence in* Y. *In particular, if* (x_n) *is an unconditional basis for* X *and* T *maps* X *onto* Y, *then* (Tx_n) *is an unconditional basis for* Y.

If (x_n) is an unconditional basis for a Banach space, then in addition to the (x_n) norm there are several other useful norms based on (x_n) and the original norm of the space that are equivalent to the original norm. The following is the most useful one for the purposes of this book. See Exercise 4.29 for another.

4.2.15 Definition. Suppose that (x_n) is an unconditional basis for a Banach space X. Then the *bounded multiplier unconditional* (x_n) *norm* of X or *bmu-*(x_n) *norm* of X is defined by the formula $\|\sum_n \alpha_n x_n\|_{\text{bmu-}(x_n)} = \sup\{\,\|\sum_n \beta_n\alpha_n x_n\| : (\beta_n) \in S_{\ell_\infty}\,\}$.

4.2.16 Theorem. *Suppose that* (x_n) *is an unconditional basis for a Banach space* X. *Then the bounded multiplier unconditional* (x_n) *norm of* X *is a Banach norm equivalent to the original norm of* X, *and* $\|x\|_{\text{bmu-}(x_n)} \geq \|x\|_{(x_n)} \geq \|x\|$ *for each* x *in* X.

PROOF. Throughout this proof, symbols denoting series convergence always indicate convergence with respect to the original norm of X, not the bmu-(x_n) norm. (It is of course a consequence of this theorem that there is really no difference between these two forms of series convergence.)

The first issue is whether $\|x\|_{\text{bmu-}(x_n)}$ is finite for each x in X. Suppose to the contrary that there is a member $\sum_n \alpha_n x_n$ of X such that $\|\sum_n \alpha_n x_n\|_{\text{bmu-}(x_n)} = +\infty$. Then for each positive integer n_1 there is a positive integer n_2 greater than n_1 and scalars $\beta_1, \dots, \beta_{n_2}$ each having absolute value no more than 1 such that

$$\left\|\sum_{n=1}^{n_2} \beta_n\alpha_n x_n\right\| \geq 1 + \sum_{n=1}^{n_1} \|\alpha_n x_n\|,$$

which implies that

$$\left\|\sum_{n=n_1+1}^{n_2} \beta_n\alpha_n x_n\right\| \geq \left\|\sum_{n=1}^{n_2} \beta_n\alpha_n x_n\right\| - \sum_{n=1}^{n_1} \|\alpha_n x_n\| \geq 1.$$

Using this fact, it is easy to construct a sequence (γ_n) in ℓ_∞ such that $\sum_n \gamma_n \alpha_n x_n$ does not converge since its sequence of partial sums is not Cauchy. This contradicts the unconditional convergence of $\sum_n \alpha_n x_n$. Thus, the function $\|\cdot\|_{\text{bmu-}(x_n)}$ is finite-valued.

It is now easy to check that $\|\cdot\|_{\text{bmu-}(x_n)}$ really is a norm. The inequalities claimed in the statement of the theorem follow immediately from Theorem 4.1.14 and the definitions of $\|\cdot\|_{\text{bmu-}(x_n)}$ and $\|\cdot\|_{(x_n)}$, which in turn implies that the identity operator from $(X, \|\cdot\|_{\text{bmu-}(x_n)})$ onto $(X, \|\cdot\|)$ is continuous. If $\|\cdot\|_{\text{bmu-}(x_n)}$ is a Banach norm, then it will be equivalent to the original norm of X by Corollary 1.6.8, so all that remains to be proved is the completeness of $(X, \|\cdot\|_{\text{bmu-}(x_n)})$.

Let $(\sum_n \alpha_{n,j} x_n)_{j=1}^\infty$ be a sequence in X that is Cauchy with respect to $\|\cdot\|_{\text{bmu-}(x_n)}$. Then this sequence is also Cauchy with respect to $\|\cdot\|$ and so converges with respect to $\|\cdot\|$ to some $\sum_n \alpha_n x_n$ having the property that $\alpha_n = \lim_j \alpha_{n,j}$ for each n. Suppose that $\epsilon > 0$. Let j_ϵ be a positive integer such that if $j, j' \geq j_\epsilon$, then $\|\sum_n \alpha_{n,j'} x_n - \sum_n \alpha_{n,j} x_n\|_{\text{bmu-}(x_n)} < \epsilon$. For each member (β_n) of S_{ℓ_∞}, each pair of integers j, j' such that $j, j' \geq j_\epsilon$, and each positive integer m,

$$\left\|\sum_{n=1}^m \beta_n \alpha_{n,j'} x_n - \sum_{n=1}^m \beta_n \alpha_{n,j} x_n\right\| \leq \left\|\sum_n \alpha_{n,j'} x_n - \sum_n \alpha_{n,j} x_n\right\|_{\text{bmu-}(x_n)} < \epsilon.$$

Letting j' tend to infinity shows that if $(\beta_n) \in S_{\ell_\infty}$ and m is a positive integer, then

$$\left\|\sum_{n=1}^m \beta_n \alpha_n x_n - \sum_{n=1}^m \beta_n \alpha_{n,j} x_n\right\| \leq \epsilon$$

whenever $j \geq j_\epsilon$. Letting m tend to infinity shows that

$$\left\|\sum_n \beta_n \alpha_n x_n - \sum_n \beta_n \alpha_{n,j} x_n\right\| \leq \epsilon$$

whenever $(\beta_n) \in S_{\ell_\infty}$ and $j \geq j_\epsilon$, from which it follows that

$$\left\|\sum_n \alpha_n x_n - \sum_n \alpha_{n,j} x_n\right\|_{\text{bmu-}(x_n)} \leq \epsilon$$

whenever $j \geq j_\epsilon$. The norm $\|\cdot\|_{\text{bmu-}(x_n)}$ is therefore a Banach norm. ■

It is possible to impose an algebra structure on a Banach space having an unconditional basis, and, if the Banach space is real, a lattice structure also. The following two order properties of bounded multiplier unconditional norms will be used in the arguments that show this.

4.2.17 Proposition. *Suppose that (x_n) is an unconditional basis for a Banach space X. Then*

$$\Big\|\sum_n \beta_n\alpha_n x_n\Big\|_{\mathrm{bmu}\text{-}(x_n)} \le \|(\beta_n)\|_\infty \Big\|\sum_n \alpha_n x_n\Big\|_{\mathrm{bmu}\text{-}(x_n)}$$

whenever $\sum_n \alpha_n x_n \in X$ and $(\beta_n) \in \ell_\infty$.

PROOF. Suppose that $\sum_n \alpha_n x_n \in X$ and $(\beta_n) \in S_{\ell_\infty}$. The proof will be finished once it is shown that $\|\sum_n \beta_n\alpha_n x_n\|_{\mathrm{bmu}\text{-}(x_n)} \le \|\sum_n \alpha_n x_n\|_{\mathrm{bmu}\text{-}(x_n)}$, and this inequality follows easily from the definition of $\|\cdot\|_{\mathrm{bmu}\text{-}(x_n)}$. ■

4.2.18 Proposition. *Suppose that (x_n) is an unconditional basis for a Banach space X. Then*

$$\Big\|\sum_n \alpha_n x_n\Big\|_{\mathrm{bmu}\text{-}(x_n)} \le \Big\|\sum_n \beta_n x_n\Big\|_{\mathrm{bmu}\text{-}(x_n)}$$

whenever $\sum_n \alpha_n x_n, \sum_n \beta_n x_n \in X$ and $|\alpha_n| \le |\beta_n|$ for each n.

PROOF. If $\sum_n \alpha_n x_n, \sum_n \beta_n x_n \in X$ and $|\alpha_n| \le |\beta_n|$ for each n, then there is a member (γ_n) of B_{ℓ_∞} such that $\alpha_n = \gamma_n\beta_n$ for each n, which by Proposition 4.2.17 implies that

$$\Big\|\sum_n \alpha_n x_n\Big\|_{\mathrm{bmu}\text{-}(x_n)} \le \|(\gamma_n)\|_\infty \Big\|\sum_n \beta_n x_n\Big\|_{\mathrm{bmu}\text{-}(x_n)} \le \Big\|\sum_n \beta_n x_n\Big\|_{\mathrm{bmu}\text{-}(x_n)},$$

as required. ■

Bounded multiplier unconditional norms provide a way to turn any Banach space having an unconditional basis into a Banach algebra. Suppose that a Banach space $(X, +, \cdot, \|\cdot\|)$ has an unconditional basis. By Corollary 4.2.13, the space has a normalized unconditional basis (x_n). Define a multiplication of the elements of X by the formula

$$\Big(\sum_n \alpha_n x_n\Big) \bullet \Big(\sum_n \beta_n x_n\Big) = \sum_n \alpha_n\beta_n x_n.$$

It follows from the bounded multiplier test that the series on the right side of this formula does converge, since the convergence of $\sum_n \alpha_n x_n$ and the fact that $\|x_n\| = 1$ for each n together imply that $(\alpha_n) \in c_0$. It is easy to check that for all x, y, z in X and every scalar α,

(1) $x \bullet (y \bullet z) = (x \bullet y) \bullet z$;

(2) $x \bullet (y + z) = (x \bullet y) + (x \bullet z)$ and $(x + y) \bullet z = (x \bullet z) + (y \bullet z)$; and

(3) $\alpha \cdot (x \bullet y) = (\alpha \cdot x) \bullet y = x \bullet (\alpha \cdot y)$.

If $\sum_n \alpha_n x_n, \sum_n \beta_n x_n \in X$, then

$$\begin{aligned}\Big\|\sum_n \alpha_n\beta_n x_n\Big\|_{\text{bmu-}(x_n)} &\le \|(\alpha_n)\|_\infty \Big\|\sum_n \beta_n x_n\Big\|_{\text{bmu-}(x_n)} \\ &\le \Big\|\sum_n \alpha_n x_n\Big\|_{\text{bmu-}(x_n)} \Big\|\sum_n \beta_n x_n\Big\|_{\text{bmu-}(x_n)}.\end{aligned}$$

Thus, whenever $x, y \in X$,

(4) $\|x \bullet y\|_{\text{bmu-}(x_n)} \le \|x\|_{\text{bmu-}(x_n)} \|y\|_{\text{bmu-}(x_n)}$.

It follows that the ordered quintuple $(X, +, \bullet, \cdot, \|\cdot\|_{\text{bmu-}(x_n)})$ is a Banach algebra. This algebra cannot have a multiplicative identity; the only candidate is $\sum_n x_n$, and that series does not converge. The algebra does, however, have the following property.

4.2.19 Definition. A Banach algebra $(X, +, \bullet, \cdot, \|\cdot\|)$ is *commutative* if $x \bullet y = y \bullet x$ whenever $x, y \in X$.

These observations can be summarized as follows.

4.2.20 Theorem. *Suppose that $(X, +, \cdot, \|\cdot\|)$ is a Banach space with an unconditional basis. Let (x_n) be a normalized unconditional basis for X, and let*

$$\Big(\sum_n \alpha_n x_n\Big) \bullet \Big(\sum_n \beta_n x_n\Big) = \sum_n \alpha_n\beta_n x_n$$

whenever $\sum_n \alpha_n x_n, \sum_n \beta_n x_n \in X$. Then $(X, +, \bullet, \cdot, \|\cdot\|_{\text{bmu-}(x_n)})$ is a commutative Banach algebra without identity.

Suppose that X is a partially ordered set and that $x, y \in X$. A *least upper bound* for x and y is an element of X, denoted by $x \vee y$, such that

(1) $x \preceq x \vee y$ and $y \preceq x \vee y$; and

(2) $x \vee y \preceq z$ whenever $x \preceq z$ and $y \preceq z$.

The appropriate modifications can be made to this to give the definition of a greatest lower bound $x \wedge y$ for x and y. Least upper bounds and greatest lower bounds are obviously unique when they exist, so the notation is unambiguous.

4.2.21 Definitions. An *ordered vector space* is a vector space X over $\mathbb{R}$ with a partial order $\preceq$ such that if $x, y \in X$ and $x \preceq y$, then

(1) $x + z \preceq y + z$ whenever $z \in X$; and

(2) $tx \preceq ty$ whenever $t > 0$.

A *vector lattice* is an ordered vector space X such that

(3) every pair of elements of X has a least upper bound.

For every element x of a vector lattice, let the *absolute value* of x be defined by the formula $|x| = x \vee (-x)$. A real normed space X that is also a vector lattice is a *normed lattice* if

(4) $\|x\| \leq \|y\|$ whenever $x, y \in X$ and $|x| \preceq |y|$.

A *Banach lattice* is a real Banach space that is a normed lattice.

Suppose that X is an ordered vector space. It is easy to check that $-y \preceq -x$ whenever $x, y \in X$ and $x \preceq y$. It readily follows that if x and y are members of a vector lattice, then $-\big((-x) \vee (-y)\big)$ is a greatest lower bound for x and y. It would therefore be redundant to require that every pair of elements of a vector lattice have a greatest lower bound; however, that requirement is sometimes inserted into the definition of a vector lattice for clarity.

4.2.22 Theorem. *Suppose that* $(X, \|\cdot\|)$ *is a real Banach space having an unconditional basis* (x_n). *Define a partial order* $\preceq$ *on* X *by declaring that* $\sum_n \alpha_n x_n \preceq \sum_n \beta_n x_n$ *when* $\alpha_n \leq \beta_n$ *for each* n. *Then* $(X, \preceq, \|\cdot\|_{\text{bmu-}(x_n)})$ *is a Banach lattice. Furthermore, if* x *and* y *are members of* X, *where* $x = \sum_n \alpha_n x_n$ *and* $y = \sum_n \beta_n x_n$, *then* $x \vee y = \sum_n \max\{\alpha_n, \beta_n\} x_n$, $x \wedge y = \sum_n \min\{\alpha_n, \beta_n\} x_n$, *and* $|x| = \sum_n |\alpha_n|\, x_n$.

PROOF. It is easy to verify that $(X, \preceq)$ is an ordered vector space. Now suppose that $x, y \in X$, where $x = \sum_n \alpha_n x_n$ and $y = \sum_n \beta_n x_n$. It follows from the bounded multiplier test that $\sum_n |\alpha_n|\, x_n$ and $\sum_n |\beta_n|\, x_n$ both converge, and therefore that $\sum_n \big(|\alpha_n| + |\beta_n|\big) x_n$ converges. The bounded multiplier test then implies that both $\sum_n \max\{\alpha_n, \beta_n\} x_n$ and $\sum_n \min\{\alpha_n, \beta_n\} x_n$ converge. It is clear that $\sum_n \max\{\alpha_n, \beta_n\} x_n$ and $\sum_n \min\{\alpha_n, \beta_n\} x_n$ are, respectively, the least upper bound and greatest lower bound for x and y, so $(X, \preceq)$ is a vector lattice. Notice that $|x| = \sum_n \max\{\alpha_n, -\alpha_n\} x_n = \sum_n |\alpha_n|\, x_n$.

Finally, suppose that $|x| \preceq |y|$, that is, that $|\alpha_n| \leq |\beta_n|$ for each n. Then $\|x\|_{\text{bmu-}(x_n)} \leq \|y\|_{\text{bmu-}(x_n)}$ by Proposition 4.2.18, so $(X, \preceq, \|\cdot\|_{\text{bmu-}(x_n)})$ is a Banach lattice. ■

4.2.23 Example. Suppose that X is c_0 or ℓ_p such that $1 \leq p < \infty$ and that (e_n) is the standard unit vector basis for X. Then the original norm of X is the same as its bmu-(e_n) norm, so the multiplication of elements of X given by the formula $(\alpha_n) \bullet (\beta_n) = (\alpha_n \beta_n)$ turns X into a commutative Banach algebra. Furthermore, if $\mathbb{F} = \mathbb{R}$, then X is a Banach lattice with respect to the partial order given by declaring that $(\alpha_n) \preceq (\beta_n)$ when $\alpha_n \leq \beta_n$ for each n.

The reader interested in Banach lattices can find out more about them from [157], while [205] is a good source of information on lattices in more general topological vector spaces.

When (x_n) is a basis for a Banach space, the (x_n) norm is a useful tool for demonstrating the continuity of the natural projections $\{ P_n : n \in \mathbb{N} \}$ for (x_n); see the proof of Theorem 4.1.15. If the basis (x_n) is unconditional, then the bmu-(x_n) norm can be used to prove the boundedness of a much larger class of linear maps.

4.2.24 Definition. Suppose that (x_n) is an unconditional basis for a Banach space X. For each subset A of $\mathbb{N}$, let P_A be the linear operator $\sum_n \alpha_n x_n \mapsto \sum_{n\in A} \alpha_n x_n$ from X into X. Then each such map is a *natural projection* for (x_n).

Notice that the maps P_n of Definition 4.1.12 are just special cases of those of Definition 4.2.24.

4.2.25 Theorem. *Suppose that (x_n) is an unconditional basis for a Banach space X. For each (β_n) in ℓ_∞, let $T_{(\beta_n)}$ be the map $\sum_n \alpha_n x_n \mapsto \sum_n \beta_n \alpha_n x_n$ from X into X. Then each such map $T_{(\beta_n)}$ is a bounded linear operator from X into X, and $\sup\{ \|T_{(\beta_n)}\| : (\beta_n) \in S_{\ell_\infty} \}$ is finite.*

PROOF. Suppose that $(\gamma_n) \in \ell_\infty$. Then $T_{(\gamma_n)}$ is clearly linear. For the proof that $T_{(\gamma_n)}$ is bounded, it may be assumed that $(\gamma_n) \in S_{\ell_\infty}$. The boundedness of $T_{(\gamma_n)}$ then follows, since

$$\Big\| T_{(\gamma_n)}\Big(\sum_n \alpha_n x_n\Big)\Big\|_{\text{bmu-}(x_n)} = \Big\|\sum_n \gamma_n \alpha_n x_n\Big\|_{\text{bmu-}(x_n)} \le \Big\|\sum_n \alpha_n x_n\Big\|_{\text{bmu-}(x_n)}$$

whenever $\sum_n \alpha_n x_n \in X$. Finally, since $\sup\{ \|T_{(\beta_n)}x\| : (\beta_n) \in S_{\ell_\infty} \} = \|x\|_{\text{bmu-}(x_n)} < +\infty$ whenever $x \in X$, the uniform boundedness principle assures that $\sup\{ \|T_{(\beta_n)}\| : (\beta_n) \in S_{\ell_\infty} \}$ is finite. ■

With all notation as in the statement of the preceding theorem, it follows that if A is any nonempty subset of S_{ℓ_∞}, then $\sup\{ \|T_{(\beta_n)}\| : (\beta_n) \in A \}$ must be finite. This fact can be used to obtain several constants important for the study of Banach spaces having unconditional bases. The two that are probably most commonly used are given in the following two corollaries, the first of which generalizes some properties of the natural projections P_n of Definition 4.1.12 proved in Section 4.1.

4.2.26 Corollary. *Suppose that (x_n) is an unconditional basis for a Banach space X. For each subset A of $\mathbb{N}$, the natural projection P_A is a bounded projection from X onto $[\{ x_n : n \in A \}]$. Furthermore, the quantity $\sup\{ \|P_A\| : A \subseteq \mathbb{N} \}$ is finite.*

PROOF. It follows immediately from the preceding theorem that each P_A is a bounded linear map and that $\sup\{\, \|P_A\| : A \subseteq \mathbb{N} \,\}$ is finite. Now fix a subset A of $\mathbb{N}$. Then P_A maps X into $[\{\, x_n : n \in A \,\}]$. Furthermore, the continuous map P_A agrees with the identity operator for X on $\langle \{\, x_n : n \in A \,\} \rangle$ and therefore on $[\{\, x_n : n \in A \,\}]$, from which it follows that P_A is a projection onto $[\{\, x_n : n \in A \,\}]$. ■

4.2.27 Corollary. *Suppose that (x_n) is an unconditional basis for a Banach space X. Let S be the collection of all sequences whose terms come from the set $\{-1, +1\}$, that is, of all sequences of signs. For each member (σ_n) of S, let $T_{(\sigma_n)}$ be the isomorphism from X onto X given by the formula $T_{(\sigma_n)}(\sum_n \alpha_n x_n) = \sum_n \sigma_n \alpha_n x_n$. Then $\sup\{\, \|T_{(\sigma_n)}\| : (\sigma_n) \in S \,\}$ is finite.*

PROOF. The only issue is whether each $T_{(\sigma_n)}$ is a vector space isomorphism from X onto X with a bounded inverse, and this is settled by noting that each $T_{(\sigma_n)}$ has itself as an inverse. ■

4.2.28 Definition. Suppose that (x_n) is an unconditional basis for a Banach space X. Let $\{\, P_A : A \subseteq \mathbb{N} \,\}$ be the collection of all natural projections for (x_n). Then $\sup\{\, \|P_A\| : A \subseteq \mathbb{N} \,\}$ is the *unconditional basis constant* for (x_n). Now let S be the collection of all sequences of signs, and let $T_{(\sigma_n)}(\sum_n \alpha_n x_n) = \sum_n \sigma_n \alpha_n x_n$ whenever $(\sigma_n) \in S$ and $\sum_n \alpha_n x_n \in X$. Then $\sup\{\, \|T_{(\sigma_n)}\| : (\sigma_n) \in S \,\}$ is the *unconditional constant* for (x_n).

The two constants of the preceding definition have characterizations analogous to those given for basis constants in Proposition 4.1.20.

4.2.29 Proposition. *Suppose that (x_n) is an unconditional basis for a Banach space X and that K_{ub} is the unconditional basis constant for (x_n). Then K_{ub} is the smallest real number M such that*

$$\Big\|\sum_{n\in A} \alpha_n x_n\Big\| \le M \Big\|\sum_n \alpha_n x_n\Big\|$$

whenever $\sum_n \alpha_n x_n \in X$ and $A \subseteq \mathbb{N}$, which is in turn the smallest real number M such that

$$\Big\|\sum_{n\in A} \alpha_n x_n\Big\| \le M \Big\|\sum_{n\in B} \alpha_n x_n\Big\|$$

for each pair A and B of finite subsets of $\mathbb{N}$ such that $A \subseteq B$ and each collection $\{\, \alpha_n : n \in B \,\}$ of scalars.

PROOF. It is a straightforward consequence of the definition of K_{ub} that

$$K_{\mathrm{ub}} = \sup\left\{ \frac{\|\sum_{n\in A} \alpha_n x_n\|}{\|\sum_n \alpha_n x_n\|} : \textstyle\sum_n \alpha_n x_n \in X \setminus \{0\},\ A \subseteq \mathbb{N} \right\},$$

from which the proposition readily follows. ■

4.2.30 Proposition. *Suppose that (x_n) is an unconditional basis for a Banach space X and that K_u is the unconditional constant for (x_n). Then K_u is the smallest real number M such that*

$$\left\|\sum_{n\in A}\alpha_n x_n - \sum_{n\in\mathbb{N}\setminus A}\alpha_n x_n\right\| \le M\left\|\sum_n \alpha_n x_n\right\|$$

whenever $\sum_n \alpha_n x_n \in X$ and $A \subseteq \mathbb{N}$, which is in turn the smallest real number M such that

$$\left\|\sum_{n\in A}\alpha_n x_n - \sum_{n\in B}\alpha_n x_n\right\| \le M\left\|\sum_{n\in A\cup B}\alpha_n x_n\right\|$$

for each pair A and B of disjoint finite subsets of $\mathbb{N}$ and each collection $\{\alpha_n : n \in A\cup B\}$ of scalars.

PROOF. Easy arguments based on the fact that

$$K_u = \sup\left\{\frac{\|\sum_{n\in A}\alpha_n x_n - \sum_{n\in\mathbb{N}\setminus A}\alpha_n x_n\|}{\|\sum_n \alpha_n x_n\|} : \sum_n \alpha_n x_n \in X\setminus\{0\},\ A\subseteq\mathbb{N}\right\}$$

prove the proposition. ■

With all notation as in the statement of the preceding proposition, the proposition essentially says that if $\sum_n \alpha_n x_n \in B_X$ and the positive integers are split into two disjoint sets A and $\mathbb{N}\setminus A$, then the two "halves" of $\sum_n \alpha_n x_n$ obtained by summing over A and $\mathbb{N}\setminus A$ separately cannot be farther than K_u units apart, and furthermore that K_u is the smallest real number having this property for every member of B_X and every subset of $\mathbb{N}$.

Since an unconditional basis has associated with it a basis constant, an unconditional basis constant, and an unconditional constant, it is worthwhile to know the relationship between these numbers.

4.2.31 Proposition. *Suppose that (x_n) is an unconditional basis for a Banach space X and that K_b, K_{ub}, and K_u are, respectively, the basis constant, unconditional basis constant, and unconditional constant for (x_n). Then $1 \le K_b \le K_{ub} \le \frac{1}{2}(1+K_u) \le K_u \le 2K_{ub}$. If X is renormed with $\|\cdot\|_{\mathrm{bmu}\text{-}(x_n)}$, then $K_u = K_{ub} = K_b = 1$.*

PROOF. It was shown in the discussion following Definition 4.1.18 that $1 \le K_b$, and it is an easy consequence of the definitions of K_b and K_{ub} that $K_b \le K_{ub}$. If $\sum_n \alpha_n x_n \in X$ and $A \subseteq \mathbb{N}$, then

$$2\left\|\sum_{n\in A}\alpha_n x_n\right\| - \left\|\sum_n \alpha_n x_n\right\| \le \left\|\sum_{n\in A}\alpha_n x_n - \sum_{n\in\mathbb{N}\setminus A}\alpha_n x_n\right\| \le K_u\left\|\sum_n \alpha_n x_n\right\|,$$

so $\|\sum_{n\in A}\alpha_n x_n\| \le \frac{1}{2}(1+K_u)\|\sum_n \alpha_n x_n\|$. It follows that $K_{ub} \le \frac{1}{2}(1+K_u)$. Now $\frac{1}{2}(1+K_u) \le K_u$ because $K_u \ge 1$, while $K_u \le 2K_{ub}$ since

$$\Big\|\sum_{n\in A}\alpha_n x_n - \sum_{n\in\mathbb{N}\setminus A}\alpha_n x_n\Big\| \le \Big\|\sum_{n\in A}\alpha_n x_n\Big\| + \Big\|\sum_{n\in\mathbb{N}\setminus A}\alpha_n x_n\Big\|$$
$$\le 2K_{ub}\Big\|\sum_n \alpha_n x_n\Big\|$$

whenever $\sum_n \alpha_n x_n \in X$ and $A \subseteq \mathbb{N}$.

Now suppose that X has been renormed with $\|\cdot\|_{\text{bmu-}(x_n)}$. A quick glance at the proof of Theorem 4.2.25 shows that, after the renorming, each of the maps $T_{(\sigma_n)}$ used in the definition of K_u has norm no more than 1. Therefore $K_u \le 1$, and so $K_u = K_{ub} = K_b = 1$. ■

The test for being a basis given in Theorem 4.1.24 has its analog for unconditional bases.

4.2.32 Theorem. *A sequence (x_n) in a Banach space X is an unconditional basis for X if and only if*

(1) *each x_n is nonzero;*

(2) *there is a real number M such that*

$$\Big\|\sum_{n\in A}\alpha_n x_n\Big\| \le M\Big\|\sum_{n\in B}\alpha_n x_n\Big\|$$

for each pair A and B of finite subsets of $\mathbb{N}$ such that $A \subseteq B$ and each collection $\{\alpha_n : n \in B\}$ of scalars; and

(3) *$[\{x_n : n \in \mathbb{N}\}] = X$.*

Proof. By results earlier in this chapter, the sequence (x_n) has properties (1), (2), and (3) if it is an unconditional basis for X.

Suppose conversely that (x_n) satisfies (1), (2), and (3). Then (x_n) is a basis for X by Theorem 4.1.24, so the only issue is whether it is unconditional. Suppose that $\sum_n \alpha_n x_n \in X$ and that $\sum_j \alpha_{n_j} x_{n_j}$ is a subseries of this series. It is enough to show that $\sum_j \alpha_{n_j} x_{n_j}$ converges. This follows immediately from the convergence of $\sum_n \alpha_n x_n$ and the fact that

$$\Big\|\sum_{j=m_1}^{m_2}\alpha_{n_j} x_{n_j}\Big\| \le M\Big\|\sum_{n=n_{m_1}}^{n_{m_2}}\alpha_n x_n\Big\|$$

whenever $m_1 \le m_2$. ■

4.2.33 Corollary. *A sequence (x_n) in a Banach space is an unconditional basic sequence if and only if each x_n is nonzero and there is a real number M*

such that

$$\Big\|\sum_{n\in A}\alpha_n x_n\Big\| \le M\Big\|\sum_{n\in B}\alpha_n x_n\Big\|$$

for each pair A and B of finite subsets of $\mathbb{N}$ such that $A\subseteq B$ and each collection $\{\,\alpha_n : n\in B\,\}$ of scalars.

4.2.34 Corollary. *Every permutation of an unconditional basic sequence in a Banach space is itself an unconditional basic sequence.*

So far, no examples of conditional bases have been given. It is time to remedy that.

4.2.35 Theorem. *The classical Schauder basis for $C[0,1]$ is a conditional basis.*

PROOF. Let $(s_n)_{n=0}^{\infty}$ be the classical Schauder basis for $C[0,1]$. For the argument about to be given, it will be helpful to refer to the graphs in Figure 4.1 on page 353. The argument will involve a subsequence of (s_n), that for convenience will be called (t_n), obtained as follows. Let $t_1 = s_2$, the member of the basis that is nonzero precisely on $(0,1)$. Let $t_2 = s_3$, the member of the basis that is nonzero precisely on $(0,1/2)$. Let $t_3 = s_6$, the member of the basis that is nonzero precisely on $(1/4,1/2)$. Let $t_4 = s_{11}$, the member of the basis that is nonzero precisely on $(1/4,3/8)$. Continue this pattern, selecting each t_n to be nonzero on an interval half as long as is the case for t_{n-1}, and arranging for the interval on which t_n is nonzero and the corresponding interval for t_{n-1} to share left endpoints if n is even and right endpoints if n is odd.

For each positive integer n, let v_n be the midpoint of the interval on which t_n is nonzero, and let $a_n = (\sum_{j=1}^{n} t_j)(v_n)$; that is, let $(v_1,a_1) = (1/2,1)$ and, when $n\ge 2$, let (v_n,a_n) be the coordinates of the new vertex added when passing from the graph of $\sum_{j=1}^{n-1} t_j$ to that of $\sum_{j=1}^{n} t_j$. Then $a_1 = 1$, $a_2 = 3/2$, and a moment's reflection on the process of passing from the graph of $\sum_{j=1}^{n-1} t_j$ to that of $\sum_{j=1}^{n} t_j$ shows that

$$a_n = 1 + \frac{a_{n-1}+a_{n-2}}{2}$$

when $n\ge 3$. Since

$$a_n - a_{n-1} = 1 - \frac{a_{n-1}-a_{n-2}}{2}$$

when $n\ge 3$, an easy induction argument shows that $1/2 \le a_n - a_{n-1} \le 3/4$ when $n\ge 2$. It follows from this that (a_n) is strictly increasing and unbounded and that $\|\sum_{j=1}^{n} t_j\|_\infty = a_n$ for each n.

Now let $b_n = \big(\sum_{j=1}^{n}(-1)^{j+1}t_j\big)(v_n)$ for each n; that is, let $(v_1,b_1) = (1/2,1)$ and, when $n\ge 2$, let (v_n,b_n) be the coordinates of the vertex added when passing from the graph of $\sum_{j=1}^{n-1}(-1)^{j+1}t_j$ to that of $\sum_{j=1}^{n}(-1)^{j+1}t_j$. Then $b_1 = 1$, $b_2 = -1/2$, and, when $n \geq 3$,

$$b_n = \begin{cases} \frac{1}{2}(b_{n-1} + b_{n-2}) + 1 & \text{if } n \text{ is odd;} \\ \frac{1}{2}(b_{n-1} + b_{n-2}) - 1 & \text{if } n \text{ is even.} \end{cases}$$

It follows by induction that $1 \leq b_n \leq 2$ when n is odd and $-1 \leq b_n \leq 0$ when n is even, so

$$\left\| \sum_{j=1}^{n} (-1)^{j+1} t_j \right\|_\infty = \max\{\, |b_j| : 1 \leq j \leq n \,\} \leq 2$$

for each n.

If (s_n) were unconditional and K_{u} were its unconditional constant, then it would have to be true that

$$a_n = \left\| \sum_{j=1}^{n} t_j \right\|_\infty \leq K_{\mathrm{u}} \left\| \sum_{j=1}^{n} (-1)^{j+1} t_j \right\|_\infty \leq 2K_{\mathrm{u}}$$

for each n, which contradicts the fact that (a_n) is an unbounded sequence of positive numbers. The basis (s_n) is therefore conditional. ■

Notice that, in the notation of the proof of the preceding theorem, the formal series $\sum_n t_n$ and $\sum_n (-1)^{n+1} t_n$ do not converge; in fact, the series $\sum_n \sigma_n t_n$ cannot converge for any sequence (σ_n) of signs since the terms of the series do not tend to 0. This illustrates the power of the methods that have been developed in this section, since the preceding proof shows that the classical Schauder basis for $C[0,1]$ is conditional without actually producing any members of $C[0,1]$ whose expansions with respect to the basis are only conditionally convergent.

It actually turns out that no basis for $C[0,1]$ is unconditional. Proofs of this can be found in [156] and [216].

4.2.36 Theorem. *The Haar basis for $L_1[0,1]$ is a conditional basis.*

PROOF. To follow this argument, it will be helpful to refer to the graphs in Figure 4.2 on page 360. Let (h_n) be the Haar basis for $L_1[0,1]$. It will be more convenient to work with the normalized version

$$(h_1, h_2, 2h_3, 2h_4, 4h_5, 4h_6, 4h_7, 4h_8, \dots)$$

of this basis; call it (h'_n). It is enough to show that (h'_n) is conditional. Let $g_1 = h'_1$, then let $g_n = h'_{2^{n-2}+1}$ when $n \geq 2$; that is, let (g_n) be the subsequence of (h'_n) formed by saving only the members of (h'_n) for which $h'_n(0) \neq 0$. It is easy to check that for each positive integer n,

$$\sum_{j=1}^{n} g_j(t) = \begin{cases} 2^{n-1} & \text{if } 0 \leq t < 2^{-(n-1)}; \\ 0 & \text{otherwise,} \end{cases}$$

so $\|\sum_{j=1}^{n} g_j\|_1 = 1$ for each n. Now form a subsequence (g_{n_j}) of (g_n) in the following way. Let $n_1 = 1$ and $n_2 = 2$; notice that $\|g_{n_1}\|_1 \geq 1/2$ and $\|g_{n_1} + g_{n_2}\|_1 \geq 2/2$. Now suppose that positive integers $n_1, \ldots, n_{p-1}$ have been chosen so that $n_1 < \cdots < n_{p-1}$ and $\|\sum_{j=1}^{k} g_{n_j}\|_1 \geq k/2$ when $1 \leq k \leq p-1$. Then there is a t_0 in $(0,1)$ such that $g_{n_{p-1}}(t_0) > 0$ and $\int_{t_0}^{1} |\sum_{j=1}^{p-1} g_{n_j}(t)|\, dt \geq (p-2)/2$. Let n_p be a positive integer such that the subset of $[0,1]$ on which g_{n_p} is nonzero lies in $[0, t_0)$. Then $n_{p-1} < n_p$, and it is easy to check that

$$\left\|\sum_{j=1}^{p} g_{n_j}\right\|_1 = \int_0^{t_0}\left|\sum_{j=1}^{p} g_{n_j}(t)\right| dt + \int_{t_0}^{1}\left|\sum_{j=1}^{p-1} g_{n_j}(t)\right| dt \geq 1 + \frac{p-2}{2} = \frac{p}{2}.$$

By induction, there is a subsequence (g_{n_j}) of (g_n) such that $\|\sum_{j=1}^{k} g_{n_j}\|_1 \geq k/2$ for each k.

Suppose that (h'_n) were an unconditional basis. Let K_{ub} be its unconditional basis constant. Then for each positive integer k,

$$\frac{k}{2} \leq \left\|\sum_{j=1}^{k} g_{n_j}\right\|_1 \leq K_{\mathrm{ub}}\left\|\sum_{j=1}^{n_k} g_j\right\|_1 = K_{\mathrm{ub}},$$

a contradiction. ■

As with the proof that the classical Schauder basis for $C[0,1]$ is conditional, the above proof does not actually produce any elements of the space whose basis expansions are only conditionally convergent.

It can be shown that $L_1[0,1]$ has no unconditional bases at all; proofs of this can be found in [156] and [216]. It does turn out that the Haar basis for $L_p[0,1]$ is unconditional if $1 < p < \infty$, but that is not as easy to show as the conditional nature of the basis when $p = 1$; see [157] and [216] for proofs. Exercise 4.30 gives an example of a Banach space that has two natural bases, one of which is unconditional and the other conditional.

Much has been said in this section about unconditional bases, but little about unconditional basic sequences. This does not mean that they are unimportant. In fact, the settling of an old problem about unconditional basic sequences sparked a burst of activity in Banach space theory in the early 1990s. Bessaga and Pełczyński, having just provided in their 1958 paper [26] one of the first published proofs that every infinite-dimensional Banach space has a basic sequence in it, did not wait long before asking whether all infinite-dimensional Banach spaces have *unconditional* basic sequences lurking in them; their paper [25] containing the question is the very next paper in the same volume of the same journal.[2] The problem

[2]Gowers and Maurey mention in [94] that although Bessaga and Pełczyński's paper apparently marked the first appearance of this question in print, Mazur was aware of it at least ten years earlier.

remained open until the summer of 1991, when Timothy Gowers found a counterexample based on earlier work of Thomas Schlumprecht [210]. Shortly thereafter, Bernard Maurey independently found essentially the same counterexample by essentially the same argument, so the two decided to publish jointly. Their paper [94] appeared in 1993. Upon seeing Gowers's and Maurey's original preprints, William B. Johnson pointed out that the proofs could be modified to show that their space was the first example of a *hereditarily indecomposable* Banach space; see [94] for the definition of this property. Acting on Johnson's observation, Gowers [92] was able to adapt the construction to produce the first example of an infinite-dimensional Banach space X having a closed subspace Y of codimension 1 not isomorphic to X, thus settling a question of Banach known as the *hyperplane problem*. Soon afterward, Gowers [91, 93] was able to produce counterexamples that settled several other major open problems in Banach space theory.

A good, brief, and accessible summary of the Schlumprecht-Gowers-Maurey-Johnson accomplishments, as well as several related problems still open at the time, can be found in Peter Casazza's 1994 book review [39].

Exercises

4.24 Let (x_n) be a sequence in a normed space X and let I be the collection of all finite subsets of $\mathbb{N}$ directed by declaring that $A \preceq B$ when $A \subseteq B$. Define a net (s_A) with index set I by letting $s_A = \sum_{n\in A} x_n$; as usual, the empty sum is defined to be the zero element of X. Prove that $\sum_n x_n$ is unconditionally convergent if and only if (s_A) converges.

4.25 Prove the following generalization of Proposition 4.2.1: *Suppose that X is a topological vector space for which X^* is a separating family. If $\sum_n x_n$ is a series in X such that $\sum_n x_{\pi(n)}$ converges for each permutation π of $\mathbb{N}$, then $\sum_n x_{\pi(n)} = \sum_n x_n$ for each permutation π of $\mathbb{N}$.* You may take it as known that rearranging the terms of an unconditionally convergent series of scalars does not change the limit of the series. (Do not be surprised if your proof is much shorter than that of Proposition 4.2.1. That proof is designed to be accessible to the reader who has just encountered series in Section 1.3 and knows nothing about dual spaces.)

4.26 Prove or disprove: *If a series $\sum_n x_n$ in a Banach space X converges unconditionally and (y_n) is a sequence in X such that $\|y_n\| \le \|x_n\|$ for each n, then $\sum_n y_n$ converges unconditionally.*

4.27 Suppose that $\sum_n x_n$ is a formal series in a Banach space. Prove that the following are equivalent.

(a) The series $\sum_n x_n$ is unconditionally convergent.

(b) For every sequence (α_n) of scalars such that $|\alpha_n| = 1$ for every n, the series $\sum_n \alpha_n x_n$ converges.

(c) For every sequence (σ_n) of signs, that is, of scalars taken from the set $\{-1, +1\}$, the series $\sum_n \sigma_n x_n$ converges.

4.28 Prove that a sequence (x_n) in a Banach space X is an unconditional basis for X if and only if $(x_{\pi(n)})$ is a basis for X for each permutation π of $\mathbb{N}$.

4.29 Suppose that (x_n) is an unconditional basis for a Banach space X. Then the norm for X defined by either the formula

$$\Big\|\sum_n \alpha_n x_n\Big\|_{\mathrm{u}\text{-}(x_n)} = \sup\Big\{\Big\|\sum_{n\in A} \alpha_n x_n\Big\| : A \subseteq \mathbb{N}\Big\}$$

or the formula

$$\Big\|\sum_n \alpha_n x_n\Big\|_{\mathrm{u}\text{-}(x_n)} = \sup\Big\{\Big\|\sum_{n\in F} \alpha_n x_n\Big\| : F \text{ is a finite subset of } \mathbb{N}\Big\}$$

is called the *unconditional* (x_n) *norm* of X.

(a) Show that the two formulas really do yield the same norm.

(b) Prove that the unconditional (x_n) norm of X is a Banach norm equivalent to the original norm of X, and that

$$\|x\|_{\mathrm{bmu}\text{-}(x_n)} \ge \|x\|_{\mathrm{u}\text{-}(x_n)} \ge \|x\|_{(x_n)} \ge \|x\|$$

for each x in X.

4.30 (a) Show that the basis for c given in Exercise 4.4 is unconditional.

(b) For each positive integer n, let x_n be the member of c whose first $n-1$ terms are 0 and the rest of whose terms are 1. Prove that (x_n) is a monotone normalized conditional basis for c. (This is called the *summing basis* for c.)

4.31 Suppose that (x_n) is an unconditional basis for a Banach space X and that K_u is the unconditional constant for (x_n).

(a) Prove that if $(\beta_n) \in \ell_\infty$ and $\sum_n \alpha_n x_n \in X$, then $\|\sum_n \beta_n\alpha_n x_n\| \le 2K_\mathrm{u}\,\|(\beta_n)\|_\infty\,\|\sum_n \alpha_n x_n\|$.

(b) Prove that if (β_n) is a member of ℓ_∞ having only real terms and $\sum_n \alpha_n x_n \in X$, then $\|\sum_n \beta_n\alpha_n x_n\| \le K_\mathrm{u}\,\|(\beta_n)\|_\infty\,\|\sum_n \alpha_n x_n\|$.

(Notice that if $(\beta_n) \in \ell_\infty$ and $\sum_n \alpha_n x_n \in X$, then there is a bounded norm-one real-linear functional u^* on X such that $u^*(\sum_n \beta_n\alpha_n x_n) = \|\sum_n \beta_n\alpha_n x_n\|$.)

4.3 Equivalent Bases

It is often useful to know when a Banach space Y has an isomorphic copy of some standard Banach space X embedded in it. If X has a basis, then it is plausible that one way to find an isomorphic copy of X inside Y would be to look for a basic sequence (y_n) in Y that is "enough like" a basis for X that $[\{y_n : n \in \mathbb{N}\}]$ must be isomorphic to X. The following definition would seem to be a good starting point when seeking a way to tell if two bases (x_n) and (y_n) for Banach spaces are enough alike that the Banach spaces have to be isomorphic.

4.3.1 Definition. Two bases (x_n) and (y_n) for Banach spaces are *equivalent* if, for every sequence (α_n) of scalars, the series $\sum_n \alpha_n x_n$ converges if and only if $\sum_n \alpha_n y_n$ converges.

The equivalence of bases really does guarantee the isomorphism of the Banach spaces they span.

4.3.2 Proposition. *Suppose that (x_n) and (y_n) are bases for the respective Banach spaces X and Y. Then (x_n) and (y_n) are equivalent if and only if there is an isomorphism T from X onto Y such that $Tx_n = y_n$ for each n.*

PROOF. It is clear that (x_n) and (y_n) are equivalent if there is an isomorphism T from X onto Y such that $Tx_n = y_n$ for each n. Suppose conversely that (x_n) and (y_n) are equivalent. If $(\sum_n \beta_{n,j} x_n)_{j=1}^\infty$ is a sequence in X that converges to some member $\sum_n \beta_n x_n$ of X and $(\sum_n \beta_{n,j} y_n)_{j=1}^\infty$ converges to some y in Y, then it follows from the continuity of the coordinate functionals for (x_n) and (y_n) that $y = \sum_n \beta_n y_n$. It is an easy consequence of the closed graph theorem and the equivalence of (x_n) and (y_n) that the map $\sum_n \alpha_n x_n \mapsto \sum_n \alpha_n y_n$ is an isomorphism from X onto Y that maps each x_n to the corresponding y_n. ∎

4.3.3 Corollary. *A basis equivalent to an unconditional basis is itself unconditional.*

One useful consequence of the preceding proposition is that a sequence that does not lie very far from a basis must itself be a basis.

4.3.4 Proposition. *Suppose that X is a Banach space, that (x_n) is a basic sequence in X, that K_b is the basis constant for (x_n), and that (y_n) is a sequence in X such that $\sum_n \|x_n\|^{-1}\|x_n - y_n\| < 1/(2K_b)$. Then (y_n) is a basic sequence equivalent to (x_n). If (x_n) is a basis for X, then so is (y_n).*

PROOF. It is easy to see that the normalized basic sequence $(\|x_n\|^{-1}x_n)$ has the same basis constant as (x_n); for example, this is a consequence of Proposition 4.1.20. It may therefore be assumed that (x_n) is normalized.

For each positive integer m, let x_m^* and P_m be, respectively, the m^{th} coordinate functional and m^{th} natural projection for (x_n), with both maps of course having domain $[\{ x_n : n \in \mathbb{N} \}]$. If $x \in [\{ x_n : n \in \mathbb{N} \}]$ and m is a positive integer greater than 1, then

$$|x_m^* x| = \|(x_m^* x)x_m\| = \|(P_m - P_{m-1})(x)\| \le 2K_b\|x\|,$$

so $\|x_m^*\| \le 2K_b$ when $m \ge 2$. An obvious modification of this argument shows that $\|x_1^*\| \le K_b$.

For each positive integer n, let z_n^* be a Hahn-Banach extension of x_n^* to X. Since $\sum_n |z_n^* x| \, \|x_n - y_n\| \le 2K_b(\sum_n \|x_n - y_n\|)\|x\|$ whenever $x \in X$, it follows that the formula $T(x) = \sum_n (z_n^* x)(x_n - y_n)$ defines a bounded linear operator from X into X and that $\|T\| \le 2K_b \sum_n \|x_n - y_n\| < 1$. Let I be the identity operator on X. By Theorem 3.3.13, the operator $I - T$ is an invertible member of the Banach algebra $B(X)$ and so is an isomorphism from X onto X. Notice that $(I - T)(x_n) = y_n$ for each n. It follows from Propositions 4.1.8 and 4.3.2 that (y_n) is a basic sequence equivalent to (x_n), and that (y_n) is a basis for X if (x_n) is. ■

4.3.5 Corollary. *Suppose that a Banach space X has a basis (x_n) and that D is a dense subset of X. Then X has a basis equivalent to (x_n) whose terms all come from D.*

The rest of this section is devoted primarily to finding characterizations of basic sequences equivalent to the standard unit vector bases for ℓ_1 and c_0 and deriving some results from those characterizations for c_0. The following two theorems give the fundamental characterizations.

4.3.6 Theorem. *Suppose that (x_n) is a sequence in a Banach space. Then (x_n) is a basic sequence equivalent to the standard unit vector basis for ℓ_1 if and only if $\sup_n \|x_n\| < +\infty$ and there is a positive constant M such that*

$$\sum_{n=1}^{m} |\alpha_n| \le M \left\| \sum_{n=1}^{m} \alpha_n x_n \right\|$$

whenever $m \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_m \in \mathbb{F}$.

PROOF. Let (e_n) be the standard unit vector basis for ℓ_1. Suppose first that (x_n) is a basic sequence equivalent to (e_n). Let T be an isomorphism from $[\{ x_n : n \in \mathbb{N} \}]$ onto ℓ_1 that maps each x_n to the corresponding e_n. Then $\sup_n \|x_n\| = \sup_n \|T^{-1} e_n\| \le \|T^{-1}\| < +\infty$. Furthermore, if $m \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_m \in \mathbb{F}$, then

$$\sum_{n=1}^{m} |\alpha_n| = \left\| \sum_{n=1}^{m} \alpha_n e_n \right\|_1 \le \|T\| \left\| \sum_{n=1}^{m} \alpha_n x_n \right\|,$$

so the claimed inequality is shown to hold by letting $M = \|T\|$.

Suppose conversely that $\sup_n \|x_n\| < +\infty$ and there is a positive constant M such that $\sum_{n=1}^m |\alpha_n| \le M \|\sum_{n=1}^m \alpha_n x_n\|$ whenever $m \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_m \in \mathbb{F}$. It must first be shown that (x_n) is a basic sequence. To this end, notice that $1 \le M\|x_n\|$ for each n, so each x_n is nonzero. Now suppose that $m_1, m_2 \in \mathbb{N}$, that $m_1 \le m_2$, and that $\alpha_1, \dots, \alpha_{m_2} \in \mathbb{F}$. Then

$$\begin{aligned}\left\|\sum_{n=1}^{m_1} \alpha_n x_n\right\| &\le \left(\sup_n \|x_n\|\right) \sum_{n=1}^{m_1} |\alpha_n| \\ &\le \left(\sup_n \|x_n\|\right) \sum_{n=1}^{m_2} |\alpha_n| \\ &\le M\left(\sup_n \|x_n\|\right) \left\|\sum_{n=1}^{m_2} \alpha_n x_n\right\|.\end{aligned}$$

It follows from Corollary 4.1.25 that (x_n) is basic.

Let (α_n) be any sequence of scalars. If $m_1, m_2 \in \mathbb{N}$ and $m_1 \le m_2$, then

$$\sum_{n=m_1}^{m_2} |\alpha_n| \le M \left\| \sum_{n=m_1}^{m_2} \alpha_n x_n \right\| \le M\left(\sup_n \|x_n\|\right) \sum_{n=m_1}^{m_2} |\alpha_n|.$$

It follows that $\sum_n \alpha_n e_n$ converges if and only if $\sum_n \alpha_n x_n$ converges, so (x_n) and (e_n) are equivalent. ■

Notice that the hypotheses of the following theorem are stronger than those of the preceding one, since the statement that (x_n) is basic is now in the hypotheses rather than in one of the equivalent statements in the conclusion. See Exercise 4.34 for the reason for this.

4.3.7 Theorem. *Suppose that (x_n) is a basic sequence in a Banach space. Then (x_n) is equivalent to the standard unit vector basis for c_0 if and only if $\inf_n \|x_n\| > 0$ and there is a positive constant M such that*

$$\left\|\sum_{n=1}^{m} \alpha_n x_n\right\| \le M \max\{\, |\alpha_n| : n = 1, \dots, m \,\}$$

whenever $m \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_m \in \mathbb{F}$.

Proof. Let (e_n) be the standard unit vector basis for c_0. Suppose first that (x_n) is equivalent to (e_n). Then there is an isomorphism T from c_0 onto $[\{\, x_n : n \in \mathbb{N} \,\}]$ that maps each e_n to the corresponding x_n. It follows that $\inf_n \|x_n\| = \inf_n \|Te_n\| > 0$. Furthermore,

$$\left\|\sum_{n=1}^{m} \alpha_n x_n\right\| \le \|T\| \left\|\sum_{n=1}^{m} \alpha_n e_n\right\|_\infty = \|T\| \max\{\, |\alpha_n| : n = 1, \dots, m \,\}$$

whenever $m \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_m \in \mathbb{F}$, so the claimed inequality is obtained by letting $M = \|T\|$.

Now suppose instead that $\inf_n \|x_n\| > 0$ and there is a positive constant M such that $\|\sum_{n=1}^m \alpha_n x_n\| \le M \max\{\, |\alpha_n| : n = 1, \dots, m \,\}$ whenever

$m \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$. Let (α_n) be a sequence of scalars. If the series $\sum_n \alpha_n x_n$ converges, then so does $\sum_n \alpha_n e_n$ since $(\alpha_n) \in c_0$. On the other hand, if $\sum_n \alpha_n e_n$ converges, then

$$\left\| \sum_{n=m_1}^{m_2} \alpha_n x_n \right\| \le M \max\{\, |\alpha_n| : n = m_1, \ldots, m_2 \,\} = M \left\| \sum_{n=m_1}^{m_2} \alpha_n e_n \right\|_\infty$$

whenever $m_1, m_2 \in \mathbb{N}$ and $m_1 \le m_2$, so $\sum_n \alpha_n x_n$ must also converge. The basic sequences (x_n) and (e_n) are therefore equivalent. ■

Another characterization of basic sequences equivalent to the standard unit vector basis for c_0 requires a brief excursion into the realm of weakly unconditionally Cauchy series.

4.3.8 Definition. A formal series $\sum_n x_n$ in a Banach space is *weakly unconditionally Cauchy* if, for each permutation π of $\mathbb{N}$, the sequence of partial sums of $\sum_n x_{\pi(n)}$ is weakly Cauchy.

4.3.9 Proposition. *Let $\sum_n x_n$ be a formal series in a Banach space X. Then the following are equivalent.*

(a) *The series $\sum_n x_n$ is weakly unconditionally Cauchy.*

(b) *For each subseries of $\sum_n x_n$, the sequence of partial sums of the subseries is weakly Cauchy.*

(c) *The series $\sum_n x^* x_n$ is absolutely convergent whenever $x^* \in X^*$.*

(d) *The series $\sum_n \alpha_n x_n$ converges whenever $(\alpha_n) \in c_0$.*

(e) *There is a positive constant M such that*

$$\left\| \sum_{n=1}^{m} \alpha_n x_n \right\| \le M \max\{\, |\alpha_n| : n = 1, \ldots, m \,\}$$

whenever $m \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$.

PROOF. For series of scalars, the properties of absolute convergence, unconditional convergence, and convergence of all subseries are equivalent. It follows that each of statements 1 through 4 below is equivalent to the next statement in the list.

1. The series $\sum_n x_n$ is weakly unconditionally Cauchy.
2. For each x^* in X^* and each permutation π of $\mathbb{N}$, the series $\sum_n x^* x_{\pi(n)}$ converges.
3. For each x^* in X^*, the series $\sum_n x^* x_n$ is absolutely convergent.
4. For each subseries $\sum_j x_{n_j}$ of $\sum_n x_n$ and each x^* in X^*, the series $\sum_j x^* x_{n_j}$ converges.

5. For each subseries of $\sum_n x_n$, the sequence of partial sums of the subseries is weakly Cauchy.

This proves the equivalence of (a), (b), and (c).

If (c) holds, then by Lemma 4.2.7 there is a nonnegative real number M such that $\sup_m \|\sum_{n=1}^m \alpha_n x_n\| \leq M\|(\alpha_n)\|_\infty$ whenever $(\alpha_n) \in \ell_\infty$, from which (e) follows. Now suppose that (e) holds and that $(\alpha_n) \in c_0$. If $m_1, m_2 \in \mathbb{N}$ and $m_1 \leq m_2$, then

$$\left\| \sum_{n=m_1}^{m_2} \alpha_n x_n \right\| \leq M \max\{\, |\alpha_n| : n = m_1, \ldots, m_2 \,\},$$

from which it follows that $\sum_n \alpha_n x_n$ converges. Therefore (e) $\Rightarrow$ (d).

The proof will be finished once it is shown that (d) $\Rightarrow$ (c). Suppose that (c) fails, that is, that there is an x_0^* in X^* such that $\sum_n |x_0^* x_n| = +\infty$. It follows that there is an increasing sequence (n_j) of positive integers such that $n_1 = 1$ and $\sum_{n=n_j}^{n_{j+1}-1} |x_0^* x_n| \geq j$ for each j. Define a member (β_n) of c_0 as follows. For each positive integer n, first let j be the positive integer such that $n_j \leq n < n_{j+1}$, then let β_n be a scalar such that $|\beta_n| = j^{-1}$ and $\beta_n x_0^* x_n = j^{-1}|x_0^* x_n|$. It follows that $\sum_n \beta_n x_0^* x_n = +\infty$, so $\sum_n \beta_n x_n$ cannot converge. Therefore (d) does not hold. ■

With the preceding proposition in hand, the promised second characterization of basic sequences equivalent to the standard unit vector basis for c_0 is now just a restatement of Theorem 4.3.7.

4.3.10 Theorem. *Suppose that (x_n) is a basic sequence in a Banach space. Then (x_n) is equivalent to the standard unit vector basis for c_0 if and only if $\inf_n \|x_n\| > 0$ and the formal series $\sum_n x_n$ is weakly unconditionally Cauchy.*

By Propositions 4.1.8 and 4.3.2, a Banach space X has in it an isomorphic copy of some particular Banach space Y with a basis (y_n) if and only if X has in it a basic sequence equivalent to (y_n). Thus, results such as Theorems 4.3.6, 4.3.7, and 4.3.10 can be useful when trying to decide whether a Banach space has certain other Banach spaces embedded in it. The next theorem is an example of such an application of Theorem 4.3.10.

4.3.11 Lemma. *Suppose that a formal series $\sum_n x_n$ in a Banach space is weakly unconditionally Cauchy but not unconditionally convergent. Then there is a subsequence (x_{n_j}) of (x_n) and an increasing sequence (m_k) of positive integers with the property that if $y_k = \sum_{j=m_k}^{m_{k+1}-1} x_{n_j}$ for each k, then (y_k) is a basic sequence equivalent to the standard unit vector basis for c_0.*

PROOF. Let (x_{n_j}) be a subsequence of (x_n) such that $\sum_j x_{n_j}$ does not converge. Then there is a positive ϵ with the property that for each positive integer N there are positive integers p_N, q_N such that $q_N \geq p_N > N$ and $\|\sum_{j=p_N}^{q_N} x_{n_j}\| \geq \epsilon$. After thinning (x_{n_j}) if necessary, it may be assumed that there is an increasing sequence (m_k) of positive integers such that $\|\sum_{j=m_k}^{m_{k+1}-1} x_{n_j}\| \geq \epsilon$ for each k. Let $y_k = \sum_{j=m_k}^{m_{k+1}-1} x_{n_j}$ for each k. Then the formal series $\sum_k y_k$ is weakly unconditionally Cauchy since $\sum_n x_n$ is so. Since $y_k \xrightarrow{w} 0$, it follows from Theorem 4.1.32 that (y_k) can be thinned to a basic sequence that will also be denoted by (y_k). By Theorem 4.3.10, the basic sequence (y_k) is equivalent to the standard unit vector basis for c_0. ■

4.3.12 Theorem. *Suppose that X is a Banach space. Then the following are equivalent.*

(a) *The space X does not have c_0 embedded in it.*

(b) *Every weakly unconditionally Cauchy series in X is convergent.*

(c) *Every weakly unconditionally Cauchy series in X is unconditionally convergent.*

PROOF. Suppose first that there is an isomorphism T from c_0 into X. Let (e_n) be the sequence of standard unit vectors of c_0. Then $\sum_n Te_n$ is weakly unconditionally Cauchy, but not convergent since $\inf_n \|Te_n\| > 0$. This shows that (b) $\Rightarrow$ (a). Since (c) obviously implies (b), all that remains to be proved is that (a) $\Rightarrow$ (c). However, this follows immediately from Lemma 4.3.11. ■

In light of the preceding theorem, it is natural to ask what conditions must be placed on a Banach space to assure that every series in the space that is in some sense "weakly unconditionally convergent" is unconditionally convergent. Before investigating this, it is necessary to decide what "weak unconditional convergence" might mean. Here are two possibilities.

4.3.13 Definition. A formal series $\sum_n x_n$ in a Banach space is *weakly reordered convergent* if $\sum_n x_{\pi(n)}$ is weakly convergent (that is, the sequence of partial sums of $\sum_n x_{\pi(n)}$ is weakly convergent) for each permutation π of $\mathbb{N}$, and is *weakly subseries convergent* if each subseries of $\sum_n x_n$ is weakly convergent.

Since norm reordered convergence and norm subseries convergence are equivalent in Banach spaces, as are weak reordered Cauchyness and weak subseries Cauchyness, it might seem likely that the two terms defined just above are equivalent. However, this turns out not to be so in any Banach space that has c_0 embedded in it; see Exercise 4.37. Though the term *weak unconditional convergence* will not be used in this book, it does occasionally appear elsewhere, and usually refers to weak reordered convergence.

It turns out that no conditions whatever need to be imposed on a Banach space to assure that each of its weakly subseries convergent series is unconditionally convergent.

4.3.14 The Orlicz-Pettis Theorem. (W. Orlicz, 1929 [177]; B. J. Pettis, 1938 [182]). *A formal series in a Banach space is weakly subseries convergent if and only if it is unconditionally convergent.*

PROOF. Every unconditionally convergent series in a Banach space has to be weakly subseries convergent since each of its subseries is actually convergent with respect to the norm topology. For the converse, suppose that some Banach space has a weakly subseries convergent series $\sum_n x_n$ in it that is not unconditionally convergent. By Lemma 4.3.11, there is a subsequence (x_{n_j}) of (x_n) and an increasing sequence (m_k) of positive integers such that, after letting $y_k = \sum_{j=m_k}^{m_{k+1}-1} x_{n_j}$ for each k, the sequence (y_n) is a basic sequence equivalent to the standard unit vector basis (e_n) for c_0. Notice that the weak convergence of $\sum_j x_{n_j}$ implies that of $\sum_n y_n$. Let T be an isomorphism from c_0 onto $[\{ y_n : n \in \mathbb{N} \}]$ such that $Te_n = y_n$ for each n. It follows from the weak-to-weak continuity of T^{-1} that $\sum_n e_n$ is weakly convergent. However, this is just not so, and this contradiction finishes the proof. ■

An immediate consequence of the Orlicz-Pettis theorem is that weak subseries convergence implies weak reordered convergence for formal series in Banach spaces. With a small amount of extra effort, it can be shown that weak subseries convergence and weak reordered convergence are equivalent notions in a Banach space X if and only if X has no copy of c_0 embedded in it; see Exercise 4.38. Therefore, a formal series in a Banach space in which c_0 is not embedded is weakly reordered convergent if and only if it is unconditionally convergent, which could be called the Orlicz-Pettis theorem for weak reordered convergence.

The proof of the Orlicz-Pettis theorem given above is due to Bessaga and Pełczyński [26]. See [58] for two other interesting proofs of the Orlicz-Pettis theorem, one due to S. Kwapień [147] that uses the Bochner integral and is based on the Pettis measurability theorem, and another that is similar to the original proofs by Orlicz and Pettis. Joseph Diestel's notes [58, pp. 29–30] on the Orlicz-Pettis theorem contain a very interesting history of this result, including the reason Pettis's name became attached to it nine years after Orlicz's original proof because of an oversight in Banach's book [13]. Other interesting discussions of this theorem can be found in [127] and [235].

The final item on this section's agenda is a look at block basic sequences, especially those taken from basic sequences equivalent to the unit vector basis for c_0 or ℓ_1.

4.3.15 Definition. Suppose that (x_n) is a basic sequence in a Banach space and that (p_n) is a sequence of positive integers such that $1 = p_1 < p_2 < p_3 < \cdots$. For each positive integer n, let $\alpha_{p_n}, \ldots, \alpha_{p_{n+1}-1}$ be scalars at least one of which is nonzero, and let $y_n = \sum_{j=p_n}^{p_{n+1}-1} \alpha_j x_j$. Then (y_n) is a *block basic sequence* taken from (x_n).

4.3.16 Proposition. *Every block basic sequence taken from a basic sequence in a Banach space is itself a basic sequence and has basis constant no more than that of the original basic sequence.*

PROOF. Suppose that (y_n) is a block basic sequence taken from a basic sequence (x_n). Then each y_n is nonzero by the uniqueness of expansions of members of $[\{x_n : n \in \mathbb{N}\}]$ in terms of (x_n). Let K be the basis constant for (x_n). Then $\|\sum_{n=1}^{m_1} \alpha_n x_n\| \leq K\|\sum_{n=1}^{m_2} \alpha_n x_n\|$ whenever $m_1, m_2 \in \mathbb{N}$, $m_1 \leq m_2$, and $\alpha_1, \ldots, \alpha_{m_2} \in \mathbb{F}$. It follows that the same is true if x_n is replaced by y_n, so (y_n) is basic by Corollary 4.1.25 and, by Proposition 4.1.20, has basis constant no more than K. ■

If a basic sequence equivalent to the standard unit vector basis for ℓ_1 or c_0 is permuted, thinned out, or blocked up, then the resulting basic sequence still generates a subspace isomorphic to ℓ_1 or c_0 respectively.

4.3.17 Theorem. *Suppose that (x_n) is a basic sequence equivalent to the standard unit vector basis for ℓ_1. Then every permutation of (x_n) and every subsequence of (x_n) is a basic sequence equivalent to the standard unit vector basis for ℓ_1. If (y_n) is a block basic sequence taken from (x_n), then $(\|y_n\|^{-1}y_n)$ is a basic sequence equivalent to the standard unit vector basis for ℓ_1, so $[\{y_n : n \in \mathbb{N}\}]$ is isomorphic to ℓ_1.*

PROOF. It follows readily from Theorem 4.3.6 that every permutation of (x_n) and every subsequence of (x_n) is a basic sequence equivalent to the standard unit vector basis (e_n) for ℓ_1.

Now suppose that (y_n) is a block basic sequence taken from (x_n). Let T be an isomorphism from ℓ_1 onto $[\{x_n : n \in \mathbb{N}\}]$ that maps each e_n to the corresponding x_n. The definition of a block basic sequence and the nature of the norm of ℓ_1 together imply that $\|\sum_{n=1}^{m} \beta_n T^{-1}y_n\|_1 = \sum_{n=1}^{m}|\beta_n|\,\|T^{-1}y_n\|_1$ whenever $m \in \mathbb{N}$ and $\beta_1, \ldots, \beta_m \in \mathbb{F}$. It follows that if $m \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$, then

$$\begin{aligned}\sum_{n=1}^{m}|\alpha_n| &\leq \|T\|\sum_{n=1}^{m}|\alpha_n|\,\|y_n\|^{-1}\|T^{-1}y_n\|_1\\ &= \|T\|\left\|\sum_{n=1}^{m}\alpha_n\|y_n\|^{-1}T^{-1}y_n\right\|_1\\ &\leq \|T\|\,\|T^{-1}\|\left\|\sum_{n=1}^{m}\alpha_n\|y_n\|^{-1}y_n\right\|.\end{aligned}$$

By Theorem 4.3.6, the basic sequence $(\|y_n\|^{-1}y_n)$ is equivalent to (e_n). ■

4.3.18 Theorem. *Suppose that (x_n) is a basic sequence equivalent to the standard unit vector basis for c_0. Then every permutation of (x_n) and* *every subsequence of* (x_n) *is a basic sequence equivalent to the standard unit vector basis for* c_0. *If* (y_n) *is a block basic sequence taken from* (x_n), *then* $(\|y_n\|^{-1}y_n)$ *is a basic sequence equivalent to the standard unit vector basis for* c_0, *so* $[\{y_n : n \in \mathbb{N}\}]$ *is isomorphic to* c_0.

PROOF. Since (x_n) is an unconditional basic sequence, each of its permutations is a basic sequence. It follows immediately from Theorem 4.3.7 that every permutation of (x_n) is a basic sequence equivalent to the standard unit vector basis (e_n) for c_0. The same theorem guarantees that every subsequence of (x_n) is equivalent to (e_n).

Now let (p_n) be a sequence of positive integers such that $1 = p_1 < p_2 < p_3 < \cdots$, and for each positive integer n let $y_n = \sum_{j=p_n}^{p_{n+1}-1} \alpha_j x_j$, where $\alpha_{p_n}, \ldots, \alpha_{p_{n+1}-1}$ are scalars that are not all zero. All that remains to be proved is that the normalization $(\|y_n\|^{-1}y_n)$ of the block basic sequence (y_n) is equivalent to (e_n). To this end, let $z_n = \sum_{j=p_n}^{p_{n+1}-1} \alpha_j e_j$ for each n. Suppose it could be proved that $(\|z_n\|_\infty^{-1}z_n)$ is equivalent to (e_n). Let T be an isomorphism from c_0 onto $[\{x_n : n \in \mathbb{N}\}]$ such that $Te_n = x_n$ for each n. It would follow from Theorem 4.3.7 that there is a positive constant M such that whenever $m \in \mathbb{N}$ and $\beta_1, \ldots, \beta_m \in \mathbb{F}$,

$$\begin{aligned}\left\|\sum_{n=1}^{m} \beta_n \|y_n\|^{-1} y_n\right\| &= \left\|T\left(\sum_{n=1}^{m} \beta_n \|Tz_n\|^{-1} z_n\right)\right\| \\ &\le \|T\|\,\|T^{-1}\| \left\|\sum_{n=1}^{m} \beta_n \|z_n\|_\infty^{-1} z_n\right\|_\infty \\ &\le M\|T\|\,\|T^{-1}\| \max\{\,|\beta_n| : n = 1, \ldots, m\,\},\end{aligned}$$

so $(\|y_n\|^{-1}y_n)$ would be equivalent to (e_n) by another application of Theorem 4.3.7. It is therefore enough to prove that $(\|z_n\|_\infty^{-1}z_n)$ is equivalent to (e_n). If $m \in \mathbb{N}$ and $\beta_1, \ldots, \beta_m \in \mathbb{F}$, then

$$\begin{aligned}&\left\|\sum_{n=1}^{m} \beta_n \|z_n\|_\infty^{-1} z_n\right\|_\infty \\ &\qquad = \left\|\sum_{n=1}^{m} \sum_{j=p_n}^{p_{n+1}-1} \beta_n (\max\{|\alpha_{p_n}|, \ldots, |\alpha_{p_{n+1}-1}|\})^{-1} \alpha_j e_j\right\|_\infty \\ &\qquad \le \max\{\,|\beta_n| : n = 1, \ldots, m\,\},\end{aligned}$$

so yet another application of Theorem 4.3.7 shows that $(\|z_n\|_\infty^{-1}z_n)$ is equivalent to (e_n). ■

The final result of this section is the following variation on the fact that every sequence in a Banach space that converges to zero weakly but not in norm has a basic subsequence; see Theorem 4.1.32.

4.3.19 The Bessaga-Pełczyński Selection Principle. (C. Bessaga and A. Pełczyński, 1958 [26]). *Suppose that (x_n^*) is the sequence of coordinate functionals for a basis (x_n) for a Banach space X and that (y_n) is a sequence in X such that $\lim_m x_n^* y_m = 0$ for each n but (y_n) does not converge to zero with respect to the norm topology. Then some subsequence of (y_n) is a basic sequence equivalent to a block basic sequence taken from (x_n).*

PROOF. Suppose first that $\|y_n\| = 1$ for each n. Let $K_{(x_n)}$ be the basis constant for (x_n) and let $q_1 = 1$. Since $\lim_m x_n^* y_m = 0$ for each n, there is a positive integer m_1 such that $\|\sum_{n=1}^{q_1}(x_n^* y_{m_1})x_n\| = \|(x_1^* y_{m_1})x_1\| < K_{(x_n)}^{-1}2^{-4}$. Since $y_{m_1} = \sum_n (x_n^* y_{m_1})x_n$, there is a positive integer q_2 such that $q_2 > q_1$ and $\|\sum_{n=q_2+1}^{\infty}(x_n^* y_{m_1})x_n\| < K_{(x_n)}^{-1}2^{-4}$. There is then a positive integer m_2 such that $m_2 > m_1$ and $\|\sum_{n=1}^{q_2}(x_n^* y_{m_2})x_n\| < K_{(x_n)}^{-1}2^{-5}$, and a positive integer q_3 such that $q_3 > q_2$ and $\|\sum_{n=q_3+1}^{\infty}(x_n^* y_{m_2})x_n\| < K_{(x_n)}^{-1}2^{-5}$. Continuing in the obvious fashion produces increasing sequences (m_n) and (q_n) of positive integers such that for each positive integer j,

$$\left\|\sum_{n=1}^{q_j}(x_n^* y_{m_j})x_n\right\| < K_{(x_n)}^{-1}2^{-(j+3)}$$

and

$$\left\|\sum_{n=q_{j+1}+1}^{\infty}(x_n^* y_{m_j})x_n\right\| < K_{(x_n)}^{-1}2^{-(j+3)}.$$

For each positive integer j, let

$$z_j = \sum_{n=q_j+1}^{q_{j+1}}(x_n^* y_{m_j})x_n.$$

It follows that for each j,

$$\begin{aligned}
1 &= \|y_{m_j}\| \\
&= \left\|\sum_n (x_n^* y_{m_j})x_n\right\| \\
&\le \left\|\sum_{n=1}^{q_j}(x_n^* y_{m_j})x_n\right\| + \|z_j\| + \left\|\sum_{n=q_{j+1}+1}^{\infty}(x_n^* y_{m_j})x_n\right\| \\
&< \|z_j\| + K_{(x_n)}^{-1}2^{-(j+2)} \\
&< \|z_j\| + 2^{-1},
\end{aligned}$$

so $\|z_j\| > 1/2 > 0$. Therefore (z_n) is a block basic sequence taken from (x_n). The basis constant $K_{(z_n)}$ for (z_n) can be no more than $K_{(x_n)}$ by Proposition 4.3.16, which together with what has been proved above shows that

$$\begin{aligned}\sum_j \|z_j\|^{-1}\|z_j - y_{m_j}\| &\le 2\sum_j \|z_j - y_{m_j}\| \\ &\le 2\sum_j \left(\left\| \sum_{n=1}^{q_j} (x_n^* y_{m_j}) x_n \right\| + \left\| \sum_{n=q_{j+1}+1}^{\infty} (x_n^* y_{m_j}) x_n \right\| \right) \\ &< 2\sum_j K_{(x_n)}^{-1} 2^{-(j+2)} \\ &= \frac{1}{2K_{(x_n)}} \\ &\le \frac{1}{2K_{(z_n)}}.\end{aligned}$$

It follows from Proposition 4.3.4 that (y_{m_n}) is a basic sequence equivalent to (z_n). This finishes the proof for the special case in which $\|y_n\| = 1$ for each n.

Now drop the assumption that each y_n has norm 1. By thinning (y_n) if necessary, it may at least be assumed that $\inf_n \|y_n\| > 0$. It follows that for each positive integer n,

$$\limsup_m \left| x_n^* \left(\|y_m\|^{-1} y_m \right) \right| \le \left(\inf_m \|y_m\| \right)^{-1} \limsup_m |x_n^* y_m| = 0,$$

and therefore that $\lim_m x_n^* \left(\|y_m\|^{-1} y_m \right) = 0$ for each n. By the special case proved above, some subsequence $\left(\|y_{m_n}\|^{-1} y_{m_n} \right)$ of $\left(\|y_n\|^{-1} y_n \right)$ is a basic sequence equivalent to some block basic sequence (z_n) taken from (x_n), so (y_{m_n}) is a basic sequence equivalent to the block basic sequence $\left(\|y_{m_n}\| z_n \right)$ taken from (x_n). ■

In this book, the only need for the Bessaga-Pełczyński selection principle will be for the proof in the next section that a Banach space has ℓ_1 embedded in it if c_0 is embedded in its dual. However, the principle has many other applications to Banach space theory. See [58] for several important ones.

Exercises

4.32 Suppose that X is $C[0,1]$ or $L_p[0,1]$ such that $1 \le p < \infty$. Show that X has a basis whose terms are all polynomials. (However, the sequence of polynomials that is the most obvious candidate to be a basis for $C[0,1]$ is not even a basic sequence. See Exercise 4.13.)

4.33 Suppose that $1 < p < \infty$. Prove the best analog of Theorems 4.3.6 and 4.3.7 for ℓ_p that you can devise.

4.34 (a) Give an example of a sequence (x_n) in a Banach space such that (x_n) is not basic even though $\inf_n \|x_n\| > 0$ and there is a positive constant M such that

$$\left\|\sum_{n=1}^{m} \alpha_n x_n\right\| \leq M \max\{\, |\alpha_n| : n = 1, \ldots, m\,\}$$

whenever $m \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$. Thus, unlike the situation for Theorem 4.3.6, the statement that (x_n) is basic must appear in the hypotheses of Theorem 4.3.7 rather than in one of the two equivalent statements in the conclusion.

(b) Give an example of a sequence (x_n) in a Banach space X and a sequence (y_n) in a Banach space Y such that for each sequence (α_n) of scalars, the formal series $\sum_n \alpha_n x_n$ converges if and only if $\sum_n \alpha_n y_n$ converges, even though (x_n) is basic and (y_n) is not.

4.35 Let (h_n) be the Haar basis for $L_1[0,1]$ and let $(x_n) = (\|h_n\|^{-1}h_n)$. Show that (x_n) is not equivalent to the standard unit vector basis (e_n) for ℓ_1, but that (x_n) does have a subsequence equivalent to (e_n).

4.36 Suppose that $\sum_n x_n$ is a formal series in a Banach space. Prove that the following are equivalent.

(a) The series $\sum_n x_n$ is weakly unconditionally Cauchy.

(b) There is a positive constant M_0 such that $\|\sum_{n\in F} \alpha_n x_n\| \leq M_0$ whenever F is a finite subset of $\mathbb{N}$ and $\{\,\alpha_n : n \in F\,\}$ is a collection of scalars each having absolute value 1. (As usual, the sum of an empty collection of members of the Banach space is defined to be the zero element of the space.)

(c) There is a positive constant M_1 such that $\|\sum_{n\in F} \sigma_n x_n\| \leq M_1$ whenever F is a finite subset of $\mathbb{N}$ and (σ_n) is a sequence of signs, that is, of scalars taken from the set $\{-1, +1\}$.

4.37 Let (e_n) be the standard unit vector basis for c_0. Let $x_1 = e_1$, and let $x_n = e_n - e_{n-1}$ when $n \geq 2$.

(a) Prove that the formal series $\sum_n x_n$ is weakly reordered convergent.

(b) Prove that the formal series $\sum_n x_n$ is not weakly subseries convergent by displaying one of its subseries that is not weakly convergent. (Notice that this also follows from the Orlicz-Pettis theorem, since the formal series does not converge.)

4.38 Suppose that X is a Banach space. Prove that the notions of weak reordered convergence and weak subseries convergence (that is, unconditional convergence; see the Orlicz-Pettis theorem) are equivalent in X if and only if X does not have c_0 embedded in it. Exercise 4.37 may help.

4.39 (a) Prove that every block basic sequence taken from a monotone basic sequence is itself monotone.

(b) Prove that every block basic sequence taken from an unconditional basic sequence is itself unconditional.

4.4 Bases and Duality

Suppose that X is c_0 or ℓ_p such that $1 < p < \infty$. Let T be the usual isometric isomorphism from ℓ_q onto X^*, where q is 1 if X is c_0 and otherwise is such that $p^{-1} + q^{-1} = 1$. Let (e_n) and (e'_n) be the standard unit vector bases for X and ℓ_q respectively, and let (e^*_n) be the sequence of coordinate functionals for (e_n). Then $Te'_n = e^*_n$ for each n, from which it follows immediately that (e^*_n) is a basis for X^*.

This leads naturally to the question of whether it is always the case that the sequence of coordinate functionals for a basis for a Banach space is a basis for the dual space, but this is quickly settled by observing that the sequence $(e^*_{1,n})$ of coordinate functionals for the standard unit vector basis $(e_{1,n})$ for ℓ_1 cannot be a basis for ℓ_1^*, since ℓ_1^* is not separable. However, a moment's thought about the natural identification of ℓ_1^* with ℓ_∞ shows that $(e^*_{1,n})$ is a basis for a subspace of ℓ_1^* isometrically isomorphic to c_0, so one could still hope that sequences of coordinate functionals for bases are themselves always basic sequences. This hope, at least, turns out not to be in vain.

4.4.1 Theorem. *Suppose that (x^*_n) is the sequence of coordinate functionals for a basis (x_n) for a Banach space X. Then (x^*_n) is a basic sequence in X^* whose basis constant is no more than that of (x_n). Furthermore, if Q is the natural map from X into X^{**}, then the sequence of coordinate functionals for (x^*_n) is formed by restricting the terms of (Qx_n) to $[\{ x^*_n : n \in \mathbb{N} \}]$.*

PROOF. Let K be the basis constant for (x_n). Suppose that m_1 and m_2 are positive integers such that $m_1 \le m_2$ and that $\alpha_1, \ldots, \alpha_{m_2}$ are scalars. Let δ be a positive number and $\sum_n \beta_n x_n$ a member of S_X such that $|(\sum_{n=1}^{m_1} \alpha_n x^*_n)(\sum_n \beta_n x_n)| \ge \|\sum_{n=1}^{m_1} \alpha_n x^*_n\| - \delta$. Then

$$\begin{aligned}
\Big\|\sum_{n=1}^{m_1} \alpha_n x^*_n\Big\| &\le \Big|\Big(\sum_{n=1}^{m_1} \alpha_n x^*_n\Big)\Big(\sum_n \beta_n x_n\Big)\Big| + \delta \\
&= \Big|\Big(\sum_{n=1}^{m_2} \alpha_n x^*_n\Big)\Big(\sum_{n=1}^{m_1} \beta_n x_n\Big)\Big| + \delta \\
&\le \Big\|\sum_{n=1}^{m_1} \beta_n x_n\Big\| \Big\|\sum_{n=1}^{m_2} \alpha_n x^*_n\Big\| + \delta \\
&\le K\Big\|\sum_n \beta_n x_n\Big\| \Big\|\sum_{n=1}^{m_2} \alpha_n x^*_n\Big\| + \delta \\
&= K\Big\|\sum_{n=1}^{m_2} \alpha_n x^*_n\Big\| + \delta.
\end{aligned}$$

Since δ is an arbitrary positive number, it follows that

$$\left\|\sum_{n=1}^{m_1} \alpha_n x_n^*\right\| \le K \left\|\sum_{n=1}^{m_2} \alpha_n x_n^*\right\|.$$

Each x_n^* is nonzero, so Corollary 4.1.25 implies that (x_n^*) is a basic sequence. By Proposition 4.1.20, the basis constant for (x_n^*) is no more than K. Finally, if $m \in \mathbb{N}$ and $\sum_n \gamma_n x_n^* \in [\{ x_n^* : n \in \mathbb{N} \}]$, then $(Qx_m)(\sum_n \gamma_n x_n^*) = \sum_n \gamma_n x_n^* x_m = \gamma_m$, so the restriction of Qx_m to $[\{ x_n^* : n \in \mathbb{N} \}]$ is the m^{th} coordinate functional for (x_n^*). ■

Consider again the example given in the second paragraph of this section. It is true that $(e_{1,n}^*)$ is not a basis for ℓ_1^* in the usual sense, but it is a basis for ℓ_1^* in a certain weak* sense. Suppose that $x^* \in \ell_1^*$ and that (α_n) is the member of ℓ_∞ that is identified with x^* in the usual way. It is easy to check that $\lim_k \sum_{n=1}^k \alpha_n e_{1,n}^* x = x^* x$ whenever $x \in \ell_1$, so $x^* = w^*\text{-}\lim_k \sum_{n=1}^k \alpha_n e_{1,n}^*$. Since $x^* e_{1,n} = \alpha_n$ for every term $e_{1,n}$ of the standard unit vector basis for ℓ_1, no sequence of scalars (β_n) besides (α_n) has the property that $x^* = w^*\text{-}\lim_k \sum_{n=1}^k \beta_n e_{1,n}^*$. Notice that $(e_{1,n})$ (or, more properly, the sequence $(Qe_{1,n})$, where Q is the natural map from ℓ_1 into ℓ_1^{**}) is the sequence of "coordinate functionals" for the "weak* basis" $(e_{1,n}^*)$ for ℓ_1^*.

As it turns out, the sequence of coordinate functionals for any Banach space basis (x_n) has this property of being a weak* basis whose sequence of coordinate functionals is (x_n) in the sense of the preceding paragraph.

4.4.2 Theorem. *Suppose that (x_n^*) is the sequence of coordinate functionals for a basis (x_n) for a Banach space X. For each x^* in X^*, there is a unique sequence (α_n) of scalars such that $x^* = w^*\text{-}\lim_k \sum_{n=1}^k \alpha_n x_n^*$. Furthermore, if Q is the natural map from X into X^{**}, then*

$$(Qx_m)\left(w^*\text{-}\lim_k \sum_{n=1}^k \alpha_n x_n^*\right) = \alpha_m$$

for each positive integer m and each sequence (α_n) of scalars such that $w^*\text{-}\lim_k \sum_{n=1}^k \alpha_n x_n^*$ exists.

PROOF. Fix an x^* in X^*. For each element $\sum_n \beta_n x_n$ of X and each positive integer k,

$$\left(\sum_{n=1}^k (x^* x_n) x_n^*\right)\left(\sum_n \beta_n x_n\right) = \sum_{n=1}^k \beta_n x^* x_n = x^*\left(\sum_{n=1}^k \beta_n x_n\right),$$

from which it follows that $\sum_{n=1}^{k}(x^*x_n)x_n^* \xrightarrow{w^*} x^*$ as $k \to \infty$. Furthermore, if (α_n) is any sequence of scalars such that $x^* = w^*\text{-}\lim_k \sum_{n=1}^{k} \alpha_n x_n^*$, then

$$x^*x_m = \lim_k \sum_{n=1}^{k} \alpha_n x_n^* x_m = \alpha_m$$

for each positive integer m, which proves the uniqueness assertion of the theorem and also shows that $(Qx_m)(w^*\text{-}\lim_k \sum_{n=1}^{k} \alpha_n x_n^*) = \alpha_m$ whenever $m \in \mathbb{N}$. ■

With all notation as in Theorem 4.4.1, the last conclusion in that theorem is, roughly speaking, that each x_m can be found in $[\{ x_n^* : n \in \mathbb{N} \}]^*$ as the m^{th} coordinate functional for the basic sequence (x_n^*). Since (x_n) is a basis for X, this leads naturally to the question of whether $[\{ x_n^* : n \in \mathbb{N} \}]^*$ might have embedded in it an entire copy of X. This turns out to be so.

4.4.3 Lemma. *Suppose that (x_n^*) is the sequence of coordinate functionals for a basis (x_n) for a Banach space X. Let K be the basis constant for (x_n) and let Q be the natural map from X into X^{**}. For each member x of X, let $\Psi(x)$ be the restriction of Qx to $[\{ x_n^* : n \in \mathbb{N} \}]$. Then Ψ is an isomorphism from X into $[\{ x_n^* : n \in \mathbb{N} \}]^*$, and $K^{-1}\|x\| \le \|\Psi x\| \le \|x\|$ for each x in X.*

PROOF. The restriction map $R\colon X^{**} \to [\{ x_n^* : n \in \mathbb{N} \}]^*$ is clearly linear, and $\|Rx^{**}\| \le \|x^{**}\|$ whenever $x^{**} \in X^{**}$. Since Q is an isometric isomorphism from X into X^{**} and $\Psi = RQ$, the map Ψ is linear and bounded, and $\|\Psi x\| \le \|x\|$ for each x in X.

Suppose that $m \in \mathbb{N}$ and that $x_0 \in \langle\{x_1, \dots, x_m\}\rangle$. Let x^* be a member of S_{X^*} such that $|x^*x_0| = \|x_0\|$, and let $y^* = \sum_{n=1}^{m}(x^*x_n)x_n^*$. Notice that

$$|y^*x_0| = \left|\sum_{n=1}^{m}(x^*x_n)(x_n^*x_0)\right| = \left|x^*\left(\sum_{n=1}^{m}(x_n^*x_0)x_n\right)\right| = |x^*x_0| = \|x_0\|.$$

If $\sum_n \alpha_n x_n \in X$, then

$$\begin{aligned}\left|y^*\left(\sum_n \alpha_n x_n\right)\right| &= \left|\sum_{n=1}^{m} \alpha_n x^* x_n\right| \\ &= \left|x^*\left(\sum_{n=1}^{m} \alpha_n x_n\right)\right| \\ &\le \left\|\sum_{n=1}^{m} \alpha_n x_n\right\| \\ &\le K\left\|\sum_n \alpha_n x_n\right\|,\end{aligned}$$

which implies that $\|y^*\| \le K$. Therefore

$$\|x_0\| = |y^*x_0| = |(\Psi x_0)(y^*)| \le \|y^*\| \, \|\Psi x_0\| \le K\|\Psi x_0\|,$$

so $K^{-1}\|x_0\| \le \|\Psi x_0\|$. Since x_0 is an arbitrary member of the dense subset $\langle\{ x_n : n \in \mathbb{N} \}\rangle$ of X, it follows that $K^{-1}\|x\| \le \|\Psi x\|$ for each x in X. Since $K^{-1}\|x\| \le \|\Psi x\| \le \|x\|$ whenever $x \in X$, the map Ψ is an isomorphism. ■

4.4.4 Proposition. *Suppose that (x_n^*) is the sequence of coordinate functionals for a basis (x_n) for a Banach space X. Then X is embedded in $[\{ x_n^* : n \in \mathbb{N} \}]^*$, and is isometrically embedded in $[\{ x_n^* : n \in \mathbb{N} \}]^*$ if (x_n) is monotone.*

PROOF. This follows immediately from the preceding lemma and the observation that the map Ψ of the lemma is an isometric isomorphism whenever $K = 1$. ■

Theorem 4.4.1 and Proposition 4.4.4 suggest several further lines of inquiry. In light of Theorem 4.4.1, it is natural to ask for a simple condition C_1 on a basis (x_n) for a Banach space X that is necessary and sufficient for the sequence (x_n^*) of coordinate functionals for (x_n) to be a basis for X^*, not merely a basic sequence. Similarly, Proposition 4.4.4 and its derivation from Lemma 4.4.3 suggest the search for a simple condition C_2 on (x_n) that is necessary and sufficient for the isomorphism Ψ of Lemma 4.4.3 to map X *onto* $[\{ x_n^* : n \in \mathbb{N} \}]^*$, which would guarantee that X is, at least up to isomorphism, a dual space. It would be particularly interesting if (x_n) were to satisfy both C_1 and C_2. In that case, the fact that (x_n) satisfies C_1 guarantees that the map Ψ of Lemma 4.4.3 is just the natural map Q from X into X^{**}, and the fact that (x_n) satisfies C_2 then assures that Q maps X onto X^{**}, that is, that X is reflexive.

The purpose of this section is to pursue these issues and to see what can be learned about the structure of Banach spaces with bases along the way. Most of the ideas and results of this section are due to R. C. James and come from his 1950 article [109].

To begin the search for a condition C_1 as described above, suppose that (x_n^*) is the sequence of coordinate functionals for a basis (x_n) for a Banach space X. A few moments' thought about the role of (x_n) as the sequence of coordinate functionals for (x_n^*) as described in Theorem 4.4.1 shows that whatever condition C_1 might be, the basis (x_n) will have it if and only if $\|x^* - \sum_{n=1}^m (x^*x_n)x_n^*\| \to 0$ as $m \to \infty$ whenever $x^* \in X^*$. Now fix an x^* in X^* and a positive integer m. For each member $\sum_n \alpha_n x_n$ of X,

$$\Big(x^* - \sum_{n=1}^m (x^*x_n)x_n^*\Big)\Big(\sum_n \alpha_n x_n\Big) = \sum_n \alpha_n x^* x_n - \sum_{n=1}^m \alpha_n x^* x_n$$
$$= x^*\Big(\sum_{n=m+1}^{\infty} \alpha_n x_n\Big),$$

so $x^* - \sum_{n=1}^{m}(x^*x_n)x_n^*$ is, roughly speaking, the part of x^* that acts on $[\{x_n : n > m\}]$. In fact, the linear functional $x^* - \sum_{n=1}^{m}(x^*x_n)x_n^*$ is what one gets by restricting x^* to $[\{x_n : n > m\}]$ and then re-extending this restriction back to a member of X^* in such a way that the re-extension is zero on $\langle\{x_1, \dots, x_m\}\rangle$. The condition C_1 should be such that (x_n) possesses it if and only if this re-extension must have small norm when m is large. This might not seem to be the same as requiring the *restriction* to have small norm when m is large, since extending a bounded linear operator to a larger space can increase its norm, but it at least suggests that the following might be the desired condition C_1. As will be seen, it is.

4.4.5 Definition. Suppose that (x_n) is a basis for a Banach space X. For each x^* in X^* and each positive integer m, let $\|x^*\|_{(m)}$ be the norm of the restriction of x^* to $[\{x_n : n > m\}]$. Then (x_n) is *shrinking* if $\lim_m \|x^*\|_{(m)} = 0$ for each x^* in X^*.

4.4.6 Example. If X is c_0 or ℓ_p such that $1 < p < \infty$ and (e_n) is the standard unit vector basis for X, then the nature of the identification of X^* with ℓ_q for some q such that $1 \le q < \infty$ assures that (e_n) is shrinking. However, the standard unit vector basis for ℓ_1 is not shrinking. Consider, for example, the member x^* of ℓ_1^* identified in the usual way with the element $(1, 1, 1, \dots)$ of ℓ_∞, and observe that $\|x^*\|_{(m)} = 1$ for each m.

4.4.7 Proposition. *Suppose that (x_n) is a basis for a Banach space X and that (x_n^*) is the sequence of coordinate functionals for (x_n). Then (x_n^*) is a basis for X^* if and only if (x_n) is shrinking.*

PROOF. Suppose first that (x_n^*) is a basis for X^*. If $\sum_n \beta_n x_n^* \in X^*$, then

$$\left\|\sum_n \beta_n x_n^*\right\|_{(m)} = \left\|\sum_{n=m+1}^{\infty} \beta_n x_n^*\right\|_{(m)} \le \left\|\sum_{n=m+1}^{\infty} \beta_n x_n^*\right\|$$

for each positive integer m, so $\lim_m \|\sum_n \beta_n x_n^*\|_{(m)} = 0$. The basis (x_n) is therefore shrinking.

Suppose conversely that (x_n) is shrinking. Let K be its basis constant and let x^* be a member of X^*. Then for each positive integer m and each member $\sum_n \alpha_n x_n$ of X,

$$\left\|\sum_{n=m+1}^{\infty} \alpha_n x_n\right\| \le \left\|\sum_n \alpha_n x_n\right\| + \left\|\sum_{n=1}^{m} \alpha_n x_n\right\| \le (1+K)\left\|\sum_n \alpha_n x_n\right\|,$$

and so

$$\left|\left(x^* - \sum_{n=1}^{m}(x^*x_n)x_n^*\right)\left(\sum_n \alpha_n x_n\right)\right| = \left|x^*\left(\sum_{n=m+1}^{\infty} \alpha_n x_n\right)\right|$$
$$\le \|x^*\|_{(m)}(1+K)\left\|\sum_n \alpha_n x_n\right\|.$$

Therefore $\|x^* - \sum_{n=1}^m (x^*x_n)x_n^*\| \le \|x^*\|_{(m)}(1+K)$ whenever $m \in \mathbb{N}$, so $\lim_m \|x^* - \sum_{n=1}^m (x^*x_n)x_n^*\| = 0$. It follows that $X^* = [\{x_n^* : n \in \mathbb{N}\}]$ and therefore that (x_n^*) is a basis for X^*. ■

With all notation as in Lemma 4.4.3, the next item on the agenda is to find a simple condition C_2 on (x_n) that is necessary and sufficient for Ψ to map X onto $[\{x_n^* : n \in \mathbb{N}\}]^*$. The search will begin with a condition that is in a sense dual to the shrinking condition.

4.4.8 Definition. A basis (x_n) for a Banach space is *boundedly complete* if, whenever a sequence (α_n) of scalars is such that $\sup_m \|\sum_{n=1}^m \alpha_n x_n\|$ is finite, the series $\sum_n \alpha_n x_n$ converges.

4.4.9 Example. Let (e_n) be the standard unit vector basis for ℓ_p, where $1 \le p < \infty$. If (α_n) is a sequence of scalars such that $\sup_m \|\sum_{n=1}^m \alpha_n e_n\|_p$ is finite, then

$$\sum_n |\alpha_n|^p = \sup_m \sum_{n=1}^m |\alpha_n|^p = \sup_m \left\|\sum_{n=1}^m \alpha_n e_n\right\|_p^p < +\infty,$$

so $\sum_n \alpha_n e_n$ converges to the element (α_n) of ℓ_p. The basis (e_n) is therefore boundedly complete. However, the standard unit vector basis $(e_{0,n})$ for c_0 is not, since $\sum_n e_{0,n}$ does not converge even though $\sup_m \|\sum_{n=1}^m e_{0,n}\|_\infty = 1$.

The shrinking property for a basis implies bounded completeness for the "dual basis."

4.4.10 Proposition. *Suppose that (x_n^*) is the sequence of coordinate functionals for a shrinking basis (x_n) for a Banach space X. Then (x_n^*) is a boundedly complete basis for X^*.*

PROOF. It follows from Proposition 4.4.7 that (x_n^*) is a basis for X^*, so the only issue is whether (x_n^*) is boundedly complete. Suppose that (α_n) is a sequence of scalars such that $(\sum_{n=1}^m \alpha_n x_n^*)_{m=1}^\infty$ is bounded. Then the Banach-Alaoglu theorem assures that $(\sum_{n=1}^m \alpha_n x_n^*)_{m=1}^\infty$ has a subnet that is weakly* convergent to some x^* in X^*. Since $\lim_m (\sum_{n=1}^m \alpha_n x_n^*)(x_k) = \alpha_k$ for each k, it follows that $x^* x_k = \alpha_k$ for each k, so

$$x^* = \sum_n (x^*x_n)x_n^* = \sum_n \alpha_n x_n^*.$$

In particular, the series $\sum_n \alpha_n x_n^*$ converges. ■

The following theorem summarizes major portions of several of the preceding results and adds the important fact that a boundedly complete "dual basic sequence" must actually be a basis for the entire dual space.

4.4.11 Theorem. *Suppose that (x_n) is a basis for a Banach space X and that (x_n^*) is the sequence of coordinate functionals for (x_n). Then the following are equivalent.*

(a) *The sequence (x_n^*) is a basis for X^*.*
(b) $[\{ x_n^* : n \in \mathbb{N} \}] = X^*$.
(c) *The basis (x_n) is shrinking.*
(d) *The basic sequence (x_n^*) is boundedly complete.*

PROOF. The implications (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) $\Rightarrow$ (d) follow immediately from Theorem 4.4.1 and Propositions 4.4.7 and 4.4.10, so it is enough to show that (d) $\Rightarrow$ (b). To this end, suppose that (x_n^*) is boundedly complete and that $x^* \in X^*$. Then $x^* = w^*\text{-}\lim_m \sum_{n=1}^m (x^*x_n)x_n^*$ by Theorem 4.4.2, so the sequence $\big(\sum_{n=1}^m (x^*x_n)x_n^*\big)_{m=1}^\infty$ is bounded by Corollary 2.6.10 and thus converges by the bounded completeness of (x_n^*). It follows immediately that $x^* = \sum_n (x^*x_n)x_n^*$, so $x^* \in [\{ x_n^* : n \in \mathbb{N} \}]$. ■

As one might have suspected from the way the plot has developed, the search for the condition C_2 that began in the discussion following Proposition 4.4.4 is about to end. Bounded completeness is that condition.

4.4.12 Lemma. *Let X, (x_n), (x_n^*), and Ψ be as in Lemma 4.4.3. Then Ψ maps X onto $[\{ x_n^* : n \in \mathbb{N} \}]^*$ if and only if (x_n) is boundedly complete.*

PROOF. Suppose first that (x_n) is boundedly complete. Let y^* be a member of $[\{ x_n^* : n \in \mathbb{N} \}]^*$. The goal is to find a member of X that Ψ maps to y^*. Let K and K' be the basis constants for (x_n) and (x_n^*) respectively. Then for each positive integer m and each member $\sum_n \beta_n x_n^*$ of $[\{ x_n^* : n \in \mathbb{N} \}]$,

$$\begin{aligned}
\left|\left(\Psi\left(\sum_{n=1}^m (y^*x_n^*)x_n\right)\right)\left(\sum_n \beta_n x_n^*\right)\right| &= \left|\left(\sum_n \beta_n x_n^*\right)\left(\sum_{n=1}^m (y^*x_n^*)x_n\right)\right| \\
&= \left|\sum_{n=1}^m (y^*x_n^*)\beta_n\right| \\
&= \left|y^*\left(\sum_{n=1}^m \beta_n x_n^*\right)\right| \\
&\le K'\|y^*\|\left\|\sum_n \beta_n x_n^*\right\|.
\end{aligned}$$

Therefore $\|\Psi(\sum_{n=1}^m (y^*x_n^*)x_n)\| \le K'\|y^*\|$ whenever $m \in \mathbb{N}$, so it follows from Lemma 4.4.3 that $\|\sum_{n=1}^m (y^*x_n^*)x_n\| \le KK'\|y^*\|$ whenever $m \in \mathbb{N}$. The bounded completeness of (x_n) then assures that $\sum_n (y^*x_n^*)x_n$ converges. Now $(\Psi(\sum_n (y^*x_n^*)x_n))(x_k^*) = y^*x_k^*$ for each k, from which it follows that $\Psi(\sum_n (y^*x_n^*)x_n) = y^*$.

Suppose conversely that Ψ maps X onto $[\{x_n^* : n \in \mathbb{N}\}]^*$. Then (Ψx_n) is a basis for $[\{x_n^* : n \in \mathbb{N}\}]^*$ and is also the sequence of coordinate functionals for (x_n^*), so (Ψx_n) is boundedly complete by Theorem 4.4.11. Since isomorphisms clearly preserve bounded completeness, the basis (x_n) is boundedly complete. ■

Notice that the first of the following two results is a partial converse of Proposition 4.4.10, while the second is an analog of Theorem 4.4.11 in which the roles of bounded completeness and the shrinking property have been exchanged.

4.4.13 Theorem. *Suppose that a Banach space X has a boundedly complete basis (x_n). Then X is isomorphic to the dual space of a Banach space with a shrinking basis, and is isometrically isomorphic to such a dual space if (x_n) is monotone.*

PROOF. Let (x_n^*) and Ψ be as in Lemma 4.4.12. Then Ψ is an isomorphism from X onto $[\{x_n^* : n \in \mathbb{N}\}]^*$, and a glance at the last paragraph of the proof of that lemma shows that the sequence of coordinate functionals for (x_n^*) is a basis for $[\{x_n^* : n \in \mathbb{N}\}]^*$. It follows from Theorem 4.4.11 that (x_n^*) is a shrinking basis for $[\{x_n^* : n \in \mathbb{N}\}]$. Furthermore, Lemma 4.4.3 assures that Ψ is an isometric isomorphism if (x_n) is monotone. ■

4.4.14 Theorem. *Suppose that (x_n) is a basis for a Banach space X and that (x_n^*) is the sequence of coordinate functionals for (x_n). Then (x_n) is boundedly complete if and only if the basic sequence (x_n^*) is shrinking.*

PROOF. It was shown in the proof of Theorem 4.4.13 that the basic sequence (x_n^*) is shrinking if (x_n) is boundedly complete. Suppose conversely that (x_n^*) is shrinking. Let Ψ be the isomorphism of Lemma 4.4.12. Then (Ψx_n) is the sequence of coordinate functionals for the shrinking basis (x_n^*) for $[\{x_n^* : n \in \mathbb{N}\}]$, and so is a boundedly complete basic sequence by Theorem 4.4.11. It follows that (x_n) is also boundedly complete. ■

The following result has already been suggested in the comments following Proposition 4.4.4.

4.4.15 Theorem. *Suppose that X is a Banach space with a basis. Then the following are equivalent.*

(a) *The space X is reflexive.*

(b) *Some basis for X is both shrinking and boundedly complete.*

(c) *Every basis for X is both shrinking and boundedly complete.*

PROOF. Suppose first that X has a basis (x_n) that is both shrinking and boundedly complete. Let (x_n^*) be the sequence of coordinate functionals for (x_n) and let Ψ be the isomorphism of Lemmas 4.4.3 and 4.4.12. Since

$[\{x_n^* : n \in \mathbb{N}\}] = X^*$ by Theorem 4.4.11, it follows directly from the definition of Ψ that Ψ is just the natural map from X into X^{**} and from Lemma 4.4.12 that $\Psi(X) = [\{x_n^* : n \in \mathbb{N}\}]^* = X^{**}$. The space X is therefore reflexive, which proves that (b) $\Rightarrow$ (a).

Since (c) obviously implies (b), all that remains to be proved is that (a) $\Rightarrow$ (c). For this, suppose that X is reflexive. Let (z_n) be a basis for X and let (z_n^*) be the sequence of coordinate functionals for (z_n). Then the closed convex subset $[\{z_n^* : n \in \mathbb{N}\}]$ of X^* is weakly closed and so is weakly* closed because the weak* and weak topologies of X^* are the same. Since $\langle\{z_n^* : n \in \mathbb{N}\}\rangle$ is weakly* dense in X^* by Theorem 4.4.2, it follows that $[\{z_n^* : n \in \mathbb{N}\}] = X^*$. Therefore (z_n) is shrinking by Theorem 4.4.11. Applying the same argument to (z_n^*) instead of (z_n) shows that (z_n^*) is also shrinking, so (z_n) is boundedly complete by Theorem 4.4.14. ■

4.4.16 Corollary. *If (x_n^*) is the sequence of coordinate functionals for a basis for a reflexive Banach space X, then (x_n^*) is a basis for X^*.*

As has already been observed, the standard unit vector basis for c_0 is not boundedly complete. It turns out that no basis for c_0 is boundedly complete, and in fact much more than that can be said.

4.4.17 Lemma. *If c_0 is embedded in the dual space of a Banach space X, then ℓ_1 is embedded in X.*

PROOF. Let $(e_{0,n})$ and $(e_{1,n})$ be the standard unit vector bases for c_0 and ℓ_1 respectively, let (e_n^*) be the basis for c_0^* that corresponds to $(e_{1,n})$ when c_0^* is identified with ℓ_1 in the usual way, let Q be the natural map from X into X^{**}, and let T be an isomorphism from c_0 into X^*. Then the adjoint T^* of T maps X^{**} onto c_0^* by Theorem 3.1.22 (b). For each positive integer n, let

$$V_n = \left\{ y^* : y^* \in c_0^*,\ |y^* e_{0,m}| < \tfrac{1}{n} \text{ when } 1 \le m < n,\ |y^* e_{0,n}| > \tfrac{1}{2} \right\},$$

a weak* neighborhood of e_n^* in c_0^*. The open mapping theorem assures that there is a positive t such that $e_n^* \in T^*(tB_{X^{**}})$ for each n, so the weak* density of $Q(tB_X)$ in $tB_{X^{**}}$ and the weak*-to-weak* continuity of T^* guarantee that for each positive integer n there is an x_n in tB_X such that $Qx_n \in (T^*)^{-1}(V_n)$; that is, such that $|(T^*Qx_n)(e_{0,n})| > \frac{1}{2}$ and $|(T^*Qx_n)(e_{0,m})| < \frac{1}{n}$ when $1 \le m < n$. Now $\lim_n (T^*Qx_n)(e_{0,m}) = 0$ for each m, but (T^*Qx_n) does not converge to 0 since $\|T^*Qx_n\| > \frac{1}{2}$ for each n, so by the Bessaga-Pełczyński selection principle there is a subsequence (x_{k_n}) of (x_n) such that $(T^*Qx_{k_n})$ is a basic sequence equivalent to some block basic sequence (z_n^*) taken from (e_n^*). Since $\big(\|z_n^*\|^{-1} z_n^*\big)$ is equivalent to $(e_{1,n})$ by Theorem 4.3.17, the basic sequence $\big(\|z_n^*\|^{-1} T^*Qx_{k_n}\big)$ must also be equivalent to $(e_{1,n})$, which implies the existence of a positive constant b such that for each positive integer n,

$$b \ge \big\| \|z_n^*\|^{-1} T^*Qx_{k_n} \big\| = \|T^*Qx_{k_n}\| \, \|z_n^*\|^{-1} > \tfrac{1}{2}\|z_n^*\|^{-1}.$$

Let S be an isomorphism from $[\{\,\|z_n^*\|^{-1}T^*Qx_{k_n} : n \in \mathbb{N}\,\}]$ onto ℓ_1 such that $S(\|z_n^*\|^{-1}T^*Qx_{k_n}) = e_{1,n}$ for each n. If $m \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$, then

$$\begin{aligned}\sum_{n=1}^m |\alpha_n| &\leq \sum_{n=1}^m 2\|z_n^*\|b|\alpha_n| \\ &= \Big\|\sum_{n=1}^m 2\|z_n^*\|b\alpha_n e_{1,n}\Big\|_1 \\ &= \Big\|\sum_{n=1}^m 2\|z_n^*\|b\alpha_n S(\|z_n^*\|^{-1}T^*Qx_{k_n})\Big\|_1 \\ &= 2b\Big\|ST^*Q\Big(\sum_{n=1}^m \alpha_n x_{k_n}\Big)\Big\|_1 \\ &\leq 2b\|S\|\,\|T^*Q\|\,\Big\|\sum_{n=1}^m \alpha_n x_{k_n}\Big\|.\end{aligned}$$

Since $\sup_n\|x_{k_n}\| \leq t$, it follows from Theorem 4.3.6 that (x_{k_n}) is a basic sequence equivalent to $(e_{1,n})$, so ℓ_1 is embedded in X. ■

4.4.18 Lemma. *If c_0 is embedded in the dual space X^* of a normed space X, then X^* is not separable.*

PROOF. By Theorem 1.11.5, it may be assumed that X is a Banach space. Let R be an isomorphism from ℓ_1 into X. Then the adjoint of R maps X^* onto ℓ_1^*, a nonseparable space. It follows from Proposition 1.12.9 (a) that X^* must itself be nonseparable. ■

With only a bit more work, it can even be shown that if a Banach space X has c_0 embedded in its dual, then there is a *complemented* copy of ℓ_1 embedded in X, and X^* has embedded in it a subspace isomorphic to ℓ_∞ and complemented by a weak*-to-weak* continuous projection. An exposition of this result of Bessaga and Pełczyński [26] can be found in Joseph Diestel's book [58, pp. 48-49].

4.4.19 Theorem. *If a Banach space has c_0 embedded in it, then no basis for the space is boundedly complete.*

PROOF. Suppose that a Banach space X has a boundedly complete basis (x_n). Let Ψ be the map of Lemmas 4.4.3 and 4.4.12. Then Ψ is an isomorphism from X onto $[\{\,x_n^* : n \in \mathbb{N}\,\}]^*$ by those two lemmas. If X were to have c_0 embedded in it, then so would the separable dual space $[\{\,x_n^* : n \in \mathbb{N}\,\}]^*$, a possibility that is ruled out by Lemma 4.4.18. ■

The preceding result has an analog for ℓ_1 and the shrinking property.

4.4.20 Theorem. *If a Banach space has ℓ_1 embedded in it, then no basis for the space is shrinking.*

PROOF. Suppose that X is a Banach space such that some isomorphism T maps ℓ_1 into X. Then the adjoint of T maps X^* onto the nonseparable space ℓ_1^*, so X^* cannot be separable and therefore cannot have a basis. By Proposition 4.4.7, the space X has no shrinking basis. ∎

The rest of this section is devoted to the shrinking and bounded completeness properties for unconditional bases. The main result for this study is the following one.

4.4.21 Theorem. *Suppose that X is a Banach space with an unconditional basis (x_n).*

(a) *The basis (x_n) is shrinking if and only if ℓ_1 is not embedded in X.*
(b) *The basis (x_n) is boundedly complete if and only if c_0 is not embedded in X.*

PROOF. Giving X a different norm equivalent to its original one does not affect whether (x_n) is shrinking or boundedly complete or whether X has ℓ_1 or c_0 embedded in it. Thus, the norm of X used in this proof will be the bmu-(x_n) norm and will be denoted by $\|\cdot\|$ rather than $\|\cdot\|_{\text{bmu-}(x_n)}$ for convenience.

Theorem 4.4.20 provides the forward implication in (a). Suppose conversely that (x_n) is not shrinking. Let x^* be an element of X^* such that, in the notation of Definition 4.4.5, the sequence $\big(\|x^*\|_{(m)}\big)$ does not tend to 0. Let $\epsilon = \frac{1}{2}\lim_m\|x^*\|_{(m)}$, a positive number. It follows that there is an increasing sequence (m_n) of positive integers and a sequence (y_n) in X such that for each positive integer n,

(1) $y_n \in \langle\{x_{m_n},\dots,x_{m_{n+1}-1}\}\rangle$;
(2) $\|y_n\| = 1$; and
(3) $x^*y_n > \epsilon$.

Suppose that $m \in \mathbb{N}$ and $\alpha_1,\dots,\alpha_m \in \mathbb{F}$. Then it follows from Proposition 4.2.18 that

$$\epsilon\sum_{n=1}^{m}|\alpha_n| \le x^*\left(\sum_{n=1}^{m}|\alpha_n|y_n\right) \le \|x^*\|\left\|\sum_{n=1}^{m}|\alpha_n|y_n\right\| = \|x^*\|\left\|\sum_{n=1}^{m}\alpha_n y_n\right\|.$$

The sequence (y_n) is therefore a basic sequence equivalent to the standard unit vector basis for ℓ_1 by Theorem 4.3.6, so ℓ_1 is embedded in X. This finishes the proof of (a).

The forward implication in (b) is an immediate consequence of Theorem 4.4.19. For the converse, suppose that (x_n) is not boundedly complete. Then there is a sequence (β_n) of scalars and a positive M such that

$\|\sum_{n=1}^m \beta_n x_n\| \le M$ for each m but $\sum_n \beta_n x_n$ does not converge, so there are sequences (p_n) and (q_n) of positive integers and a positive δ such that $p_1 \le q_1 < p_2 \le q_2 < p_3 \le q_3 < \cdots$ and $\|\sum_{j=p_n}^{q_n} \beta_j x_j\| \ge \delta$ whenever $n \in \mathbb{N}$. Let $z_n = \sum_{j=p_n}^{q_n} \beta_j x_j$ for each n. Then (z_n) is a block basic sequence taken from (x_n). It follows from Proposition 4.2.18 that if $m \in \mathbb{N}$ and $\alpha_1, \dots, \alpha_m \in \mathbb{F}$, then

$$\begin{aligned}\left\|\sum_{n=1}^m \alpha_n z_n\right\| &\le \left\|\sum_{n=1}^m \max\{\,|\alpha_n| : n = 1, \dots, m\,\} z_n\right\| \\ &= \max\{\,|\alpha_n| : n = 1, \dots, m\,\}\left\|\sum_{n=1}^m z_n\right\| \\ &\le \max\{\,|\alpha_n| : n = 1, \dots, m\,\}\left\|\sum_{n=1}^{q_m} \beta_n x_n\right\| \\ &\le M \max\{\,|\alpha_n| : n = 1, \dots, m\,\}.\end{aligned}$$

Therefore (z_n) is equivalent to the standard unit vector basis for c_0 by Theorem 4.3.7, and so c_0 is embedded in X. ■

4.4.22 Corollary. *Suppose that X is a Banach space.*

(a) *Either every unconditional basis for X is shrinking, or none is.*

(b) *Either every unconditional basis for X is boundedly complete, or none is.*

4.4.23 Corollary. *Suppose that X is a Banach space with an unconditional basis. Then X is reflexive if and only if neither c_0 nor ℓ_1 is embedded in X.*

PROOF. This follows immediately from Theorems 4.4.21 and 4.4.15. ■

4.4.24 Corollary. *Suppose that X is a Banach space in which there is an unconditional basic sequence. Then X has an infinite-dimensional reflexive subspace, or a subspace isomorphic to c_0, or a subspace isomorphic to ℓ_1.*

The preceding corollary points out the special importance of Bessaga and Pełczyński's question of whether all infinite-dimensional Banach spaces have unconditional basic sequences in them; see the discussion at the end of Section 4.2. As long as that question remained open, there was at least a small amount of hope that every infinite-dimensional Banach space might have an infinite-dimensional subspace that is reflexive or isomorphic to either c_0 or ℓ_1. That hope faded rapidly when Gowers and Maurey produced their example of an infinite-dimensional Banach space with no unconditional basic sequence in it, and vanished altogether when Gowers published in a 1994 paper [91] an example of an infinite-dimensional Banach space with no infinite-dimensional reflexive subspace and in which neither c_0 nor ℓ_1 is embedded.

See Exercises 4.41 and 4.42 for two other applications of the ideas explored in the last few results.

Exercises

4.40 Prove that $C[0,1]$ is not isomorphic to the dual space of any normed space.

4.41 Suppose that X is a Banach space with an unconditional basis. Prove that if X^* is separable, then X is reflexive if and only if c_0 is not embedded in X.

4.42 Suppose that X is a Banach space with an unconditional basis. Prove that X^{**} is separable if and only if X is reflexive. Exercise 4.41 may help.

*4.5 James's Space J

None of the results of this section depend on those of any other optional section.

The purpose of this section is to continue the development of some of the ideas of the preceding section, with the goal of being able to examine some of the unusual characteristics of an important Banach space J first devised and studied by R. C. James. Many of this space's important properties are related to the fact that it has a natural monotone shrinking basis, so the first portion of this section will focus on properties of such bases and spaces that have them.

4.5.1 Lemma. *Suppose that (x_n) is a basis for a Banach space. Let*

$$S = \left\{ (\alpha_n) : (\alpha_n) \textit{ is a sequence of scalars, } \sup_m \left\| \sum_{n=1}^{m} \alpha_n x_n \right\| < +\infty \right\}.$$

Let $\|(\alpha_n)\|_S = \sup_m \|\sum_{n=1}^m \alpha_n x_n\|$ for each member (α_n) of S. Then S, with the usual vector space operations for spaces of sequences of scalars and the norm $\|\cdot\|_S$, is a Banach space.

PROOF. It is easy to check that S is a vector space and just as easy to check that $\|\cdot\|_S$ is a norm for S. The only remaining issue is whether $\|\cdot\|_S$ is a Banach norm. Suppose that $\big((\alpha_{n,j})\big)_{j=1}^\infty$ is a Cauchy sequence in S. If $m, j, j' \in \mathbb{N}$ and $m \geq 2$, then

$$\begin{aligned} |\alpha_{m,j} - \alpha_{m,j'}| &= \|x_m\|^{-1} \left\| \sum_{n=1}^{m} (\alpha_{n,j} - \alpha_{n,j'}) x_n - \sum_{n=1}^{m-1} (\alpha_{n,j} - \alpha_{n,j'}) x_n \right\| \\ &\leq 2\|x_m\|^{-1} \|(\alpha_{n,j}) - (\alpha_{n,j'})\|_S, \end{aligned}$$

while

$$|\alpha_{1,j} - \alpha_{1,j'}| = \|x_1\|^{-1}\|(\alpha_{1,j} - \alpha_{1,j'})x_1\| \le \|x_1\|^{-1}\|(\alpha_{n,j}) - (\alpha_{n,j'})\|_S.$$

It follows that for each positive integer m there is a scalar α_m such that $\lim_j \alpha_{m,j} = \alpha_m$.

Suppose that $\epsilon > 0$. Let j_ϵ be a positive integer such that if $j, j' \ge j_\epsilon$, then $\|(\alpha_{n,j}) - (\alpha_{n,j'})\|_S < \epsilon$. If $m \in \mathbb{N}$ and $j, j' \ge j_\epsilon$, then

$$\left\|\sum_{n=1}^{m}(\alpha_{n,j} - \alpha_{n,j'})x_n\right\| \le \|(\alpha_{n,j}) - (\alpha_{n,j'})\|_S < \epsilon.$$

Letting j' tend to infinity shows that

$$\left\|\sum_{n=1}^{m}(\alpha_{n,j} - \alpha_n)x_n\right\| \le \epsilon$$

whenever $m \in \mathbb{N}$ and $j \ge j_\epsilon$, and therefore that

$$\sup_m\left\|\sum_{n=1}^{m}(\alpha_{n,j} - \alpha_n)x_n\right\| \le \epsilon$$

whenever $j \ge j_\epsilon$. It follows that $(\alpha_n) \in S$ and that $\|(\alpha_{n,j}) - (\alpha_n)\|_S \to 0$ as $j \to \infty$, so Cauchy sequences in S do converge. ■

4.5.2 Proposition. *Suppose that (x_n) is a shrinking basis for a Banach space X. Let S be the Banach space of the preceding lemma. Let (x_n^*) be the sequence of coordinate functionals for (x_n) and let $\phi(x^{**}) = (x^{**}x_n^*)$ whenever $x^{**} \in X^{**}$. Then ϕ is an isomorphism from X^{**} onto S. If (x_n) is monotone, then ϕ is an isometric isomorphism, and so $\|x^{**}\| = \lim_m\|\sum_{n=1}^{m}(x^{**}x_n^*)x_n\|$ whenever $x^{**} \in X^{**}$.*

Proof. It may be assumed that the norm of X is its (x_n) norm and therefore that (x_n) is monotone. Then (x_n^*) is a monotone basis for X^* by Theorem 4.4.1 and Proposition 4.4.7. For each element x^{**} of X^{**}, each x^* in X^*, and each positive integer m,

$$\begin{aligned}\left|x^*\left(\sum_{n=1}^{m}(x^{**}x_n^*)x_n\right)\right| &= \left|x^{**}\left(\sum_{n=1}^{m}(x^*x_n)x_n^*\right)\right|\\ &\le \|x^{**}\|\left\|\sum_{n=1}^{m}(x^*x_n)x_n^*\right\|\\ &\le \|x^{**}\|\,\|x^*\|\end{aligned}$$

since $x^* = \sum_n(x^*x_n)x_n^*$. Therefore $\|\sum_{n=1}^{m}(x^{**}x_n^*)x_n\| \le \|x^{**}\|$ whenever $x^{**} \in X^{**}$ and $m \in \mathbb{N}$, and so $\phi(x^{**}) = (x^{**}x_n^*) \in S$ whenever

$x^{**} \in X^{**}$. Notice that $\|\phi x^{**}\|_S = \sup_m \|\sum_{n=1}^m (x^{**}x_n^*)x_n\| \leq \|x^{**}\|$ for each x^{**} in X^{**}.

It is clear that ϕ is linear. If $x^{**} \in X^{**}$ and $\phi x^{**} = 0$, then for each x^* in X^*,

$$x^{**}x^* = x^{**}\left(\sum_n (x^*x_n)x_n^*\right) = \sum_n (x^*x_n)(x^{**}x_n^*) = 0,$$

so $x^{**} = 0$. The map ϕ is therefore one-to-one.

Suppose that $(\alpha_n) \in S$. If $x^* \in X^*$, the positive integers m_1 and m_2 are such that $m_1 < m_2$, and $\|\cdot\|_{(m)}$ has the meaning given in Definition 4.4.5, then

$$\left|\sum_{n=m_1+1}^{m_2} \alpha_n x^* x_n\right| \leq \|x^*\|_{(m_1)} \left\|\sum_{n=m_1+1}^{m_2} \alpha_n x_n\right\| \leq 2\|x^*\|_{(m_1)}\, \|(\alpha_n)\|_S,$$

which implies that $\sum_n \alpha_n x^* x_n$ converges since $\lim_m \|x^*\|_{(m)} = 0$. Define $x_0^{**}\colon X^* \to \mathbb{F}$ by the formula $x_0^{**}(x^*) = \sum_n \alpha_n x^* x_n$. Then x_0^{**} is clearly linear. If $x^* \in X^*$ and $m \in \mathbb{N}$, then

$$\left|\sum_{n=1}^{m} \alpha_n x^* x_n\right| \leq \|x^*\| \left\|\sum_{n=1}^{m} \alpha_n x_n\right\| \leq \|x^*\|\, \|(\alpha_n)\|_S,$$

so $|x_0^{**}x^*| \leq \|(\alpha_n)\|_S \, \|x^*\|$ whenever $x^* \in X^*$. It follows that $x_0^{**} \in X^{**}$ and $\phi x_0^{**} = (x_0^{**}x_n^*) = (\alpha_n)$, so ϕ maps X^{**} onto S.

Notice that the argument just given, along with the fact that ϕ is one-to-one, shows that $|x^{**}x^*| \leq \|\phi x^{**}\|_S \, \|x^*\|$ whenever $x^{**} \in X^{**}$ and $x^* \in X^*$, and therefore that $\|x^{**}\| \leq \|\phi x^{**}\|_S$ whenever $x^{**} \in X^{**}$. The reverse inequality was obtained earlier, so ϕ is an isometric isomorphism from X^{**} onto S. The fact that (x_n) is monotone then assures that

$$\|x^{**}\| = \|\phi x^{**}\|_S = \sup_m \left\|\sum_{n=1}^{m} (x^{**}x_n^*)x_n\right\| = \lim_m \left\|\sum_{n=1}^{m} (x^{**}x_n^*)x_n\right\|$$

whenever $x^{**} \in X^{**}$. ■

Let X, (x_n), (x_n^*), and S be as in the preceding proposition, with the stipulation that (x_n) is a *monotone* shrinking basis for X. For the purposes of this section, the main point of the proposition is that X^{**} can be identified in a natural way with the space of all sequences (α_n) of scalars such that $\sup_m \|\sum_{n=1}^m \alpha_n x_n\|$ is finite. Since (x_n^*) is a basis for X^* by Proposition 4.4.7, it follows that X^* can be identified with the space of all sequences (α_n) of scalars such that $\sum_n \alpha_n x_n^*$ converges, while of course X itself can be identified with the space of all sequences (α_n) of scalars such that $\sum_n \alpha_n x_n$ converges; see Proposition 4.1.7. To summarize, the space X

Space	Identified with the space of sequences of scalars (α_n) such that	Norm
X	$\left(\sum_{n=1}^{m} \alpha_n x_n\right)_{m=1}^{\infty}$ converges	$\|(\alpha_n)\| = \lim_m \left\|\sum_{n=1}^{m} \alpha_n x_n\right\| = \sup_m \left\|\sum_{n=1}^{m} \alpha_n x_n\right\|$
X^*	$\left(\sum_{n=1}^{m} \alpha_n x_n^*\right)_{m=1}^{\infty}$ converges	$\|(\alpha_n)\| = \lim_m \left\|\sum_{n=1}^{m} \alpha_n x_n^*\right\| = \sup_m \left\|\sum_{n=1}^{m} \alpha_n x_n^*\right\|$
X^{**}	$\left(\sum_{n=1}^{m} \alpha_n x_n\right)_{m=1}^{\infty}$ is bounded	$\|(\alpha_n)\| = \lim_m \left\|\sum_{n=1}^{m} \alpha_n x_n\right\| = \sup_m \left\|\sum_{n=1}^{m} \alpha_n x_n\right\|$

TABLE 4.1. The natural identification of X, X^*, and X^{**} with sequence spaces when X is a Banach space with a monotone shrinking basis (x_n) and (x_n^*) is the sequence of coordinate functionals for (x_n).

and its first and second dual spaces can be treated as if they were sequence spaces as in Table 4.1, where the equalities of limits and suprema in the norm column are due to the fact that (x_n) and (x_n^*) are both monotone, the first by hypothesis and the second by Theorem 4.4.1.

Suppose that the sequences of scalars (α_n), (β_n), and (γ_n) have been identified with the members x^{**}, x^*, and x of X^{**}, X^*, and X respectively. Then the actions of (α_n) on (β_n) and of (β_n) on (γ_n) are given by the formulas

$$(\alpha_n)(\beta_n) = \sum_n \alpha_n \beta_n$$

and

$$(\beta_n)(\gamma_n) = \sum_n \beta_n \gamma_n$$

since

$$x^{**}x^* = \sum_n \beta_n x^{**} x_n^* = \sum_n \beta_n \alpha_n$$

and

$$x^*x = \Big(\sum_n \beta_n x_n^*\Big)\Big(\sum_n \gamma_n x_n\Big) = \sum_n \beta_n\gamma_n.$$

Furthermore, if Q is the natural map from X into X^{**} and the identifications of Table 4.1 are made, then it is clear from the formulas for $(\alpha_n)(\beta_n)$ and $(\beta_n)(\gamma_n)$ given above that $Q(\gamma_n) = (\gamma_n)$ for each member (γ_n) of X.

Now suppose that X is c_0 and that (x_n) is its standard unit vector basis, which is indeed monotone and shrinking. Let W, Y, and Z be the sequence spaces that the correspondences of Table 4.1 identify with c_0, c_0^*, and c_0^{**} respectively. It is easy to check that W and Z are just c_0 and ℓ_∞, respectively, with their usual norms, and an easy argument based on the standard way that c_0^* is identified with ℓ_1 shows that Y is just ℓ_1 with its usual norm. Furthermore, the identification of c_0 with itself given in the table is just the identity map, while the identifications of c_0^* and c_0^{**} with ℓ_1 and ℓ_∞ respectively that follow from the table are precisely the standard identifications of these spaces that have been discussed previously in this book.[3] Thus, the identifications of Table 4.1 are generalizations of the usual characterizations of the first and second duals of c_0 to other Banach spaces having monotone shrinking bases.

The identifications described in Table 4.1 will prove useful for analyzing the space about to be defined.

4.5.3 Definition. For each member (α_n) of the vector space V underlying real c_0, let

$$\|(\alpha_n)\|_a = 2^{-1/2}\sup\Big\{ \Big(\sum_{n=1}^{m-1}(\alpha_{p_n} - \alpha_{p_{n+1}})^2 + (\alpha_{p_m} - \alpha_{p_1})^2\Big)^{1/2} : m \ge 2,\, p_1 < \cdots < p_m \Big\}$$

and

$$\|(\alpha_n)\|_b = 2^{-1/2}\sup\Big\{ \Big(\sum_{n=1}^{m-1}(\alpha_{p_n} - \alpha_{p_{n+1}})^2\Big)^{1/2} : m \ge 2,\, p_1 < \cdots < p_m \Big\}.$$

Then *James's space* J is the real Banach space of all members (α_n) of V such that $\|(\alpha_n)\|_a$ is finite, with the norm $\|\cdot\|_a$.

Though the function $\|\cdot\|_b$ is not actually used to define J, this function will have important uses in the study of the space to be done below. Of

[3] See Example 1.11.2 for the standard identification of c_0^{**} with ℓ_∞. It is easy to check directly that if that identification has been made and ℓ_1 has been identified with c_0^* in the natural way, then $(\alpha_n)(\beta_n) = \sum_n \alpha_n\beta_n$ whenever $(\alpha_n) \in c_0^{**}$ and $(\beta_n) \in c_0^*$.

course, the first issue in the analysis of J is whether it truly is a Banach space.

4.5.4 Proposition. *The space J is a real Banach space, and $\|\cdot\|_b$ is a norm for this space equivalent to $\|\cdot\|_a$.*

PROOF. The triangle inequality for $\|\cdot\|_a$ follows readily from that for ℓ_2, and it is apparent that $\|\beta(\alpha_n)\|_a = |\beta|\,\|(\alpha_n)\|_a$ whenever $(\alpha_n) \in J$ and $\beta \in \mathbb{R}$. It follows that J is a vector subspace of the vector space that underlies real c_0. It is apparent that $\|0\|_a = 0$ and that $\|(\alpha_n)\|_a > 0$ whenever (α_n) is a nonzero member of J, so $\|\cdot\|_a$ is a norm for J. It is just as easy to check that $\|\cdot\|_b$ is also a norm for J, and that this norm is equivalent to $\|\cdot\|_a$ since $\|(\alpha_n)\|_b \le \|(\alpha_n)\|_a \le 2\|(\alpha_n)\|_b$ whenever $(\alpha_n) \in J$; this last inequality uses the fact that for each member (α_n) of J,

$$\begin{aligned}\Big(\sum_{n=1}^{m-1}(\alpha_{p_n} - \alpha_{p_{n+1}})^2 + (\alpha_{p_m} - \alpha_{p_1})^2\Big)^{1/2} &\le \Big(\sum_{n=1}^{m-1}(\alpha_{p_n} - \alpha_{p_{n+1}})^2\Big)^{1/2} + \big((\alpha_{p_m} - \alpha_{p_1})^2\big)^{1/2}\\ &\le 2\cdot 2^{1/2}\|(\alpha_n)\|_b\end{aligned}$$

whenever $m \ge 2$ and $p_1 < \cdots < p_m$.

All that remains to be proved is that $\|\cdot\|_a$ is a Banach norm. Suppose that $\big((\beta_{n,j})\big)_{j=1}^{\infty}$ is a Cauchy sequence of elements of J. If $(\alpha_n) \in J$ and $n_0, n_1 \in \mathbb{N}$, then

$$\begin{aligned}|\alpha_{n_0} - \alpha_{n_1}| &= 2^{-1/2}\big((\alpha_{n_0} - \alpha_{n_1})^2 + (\alpha_{n_1} - \alpha_{n_0})^2\big)^{1/2}\\ &\le \|(\alpha_n)\|_a.\end{aligned}$$

Letting n_1 tend to ∞ shows that $|\alpha_{n_0}| \le \|(\alpha_n)\|_a$ whenever $(\alpha_n) \in J$ and $n_0 \in \mathbb{N}$. From the fact that $\big((\beta_{n,j})\big)_{j=1}^{\infty}$ is Cauchy, it follows that $(\beta_{n,j})_{j=1}^{\infty}$ is a Cauchy, therefore convergent, sequence of real numbers whenever $n \in \mathbb{N}$. Let $\beta_n = \lim_j \beta_{n,j}$ for each n.

Let ϵ be an arbitrary positive number and let j_ϵ be a positive integer such that $\|(\beta_{n,j}) - (\beta_{n,j'})\|_a < \epsilon$ when $j, j' \ge j_\epsilon$. If $j, j', m, p_1, \dots, p_m$ are positive integers such that $j, j' \ge j_\epsilon$, $m \ge 2$, and $p_1 < \cdots < p_m$, then

$$\begin{aligned}2^{-1/2}\Big(\sum_{n=1}^{m-1}\big((\beta_{p_n,j}-\beta_{p_n,j'}) - (\beta_{p_{n+1},j} - \beta_{p_{n+1},j'})\big)^2 \\ + \big((\beta_{p_m,j} - \beta_{p_m,j'}) - (\beta_{p_1,j} - \beta_{p_1,j'})\big)^2\Big)^{1/2}\\ \le \|(\beta_{n,j}) - (\beta_{n,j'})\|_a\\ < \epsilon.\end{aligned}$$

Letting j' tend to ∞ shows that

$$2^{-1/2}\left(\sum_{n=1}^{m-1}\left((\beta_{p_n,j}-\beta_{p_n})-(\beta_{p_{n+1},j}-\beta_{p_{n+1}})\right)^2 + \left((\beta_{p_m,j}-\beta_{p_m})-(\beta_{p_1,j}-\beta_{p_1})\right)^2\right)^{1/2} \le \epsilon$$

whenever $j \ge j_\epsilon$, $m \ge 2$, and $p_1 < \cdots < p_m$. It follows that $(\beta_n) \in J$ and that

$$\|(\beta_{n,j}) - (\beta_n)\|_a \le \epsilon$$

whenever $j \ge j_\epsilon$. Since $(\beta_{n,j}) \to (\beta_n)$ as $j \to \infty$, the space J is complete. ■

Let (e_n) be the sequence of standard unit vectors in real c_0. Then each of these vectors is in J, and $\|e_n\|_a = 1$ for each n. The sequence (e_n) turns out to be a monotone shrinking basis for J just as it is for c_0.

4.5.5 Proposition. *The sequence (e_n) of standard unit vectors in J is a monotone shrinking basis for J, and $(\alpha_n) = \sum_n \alpha_n e_n$ whenever $(\alpha_n) \in J$.*

PROOF. If $(\alpha_n) \in J$, then the sequence $\left(\|(\alpha_1,\dots,\alpha_k,0,0,\dots)\|_a\right)_{k=1}^\infty$ is clearly nondecreasing, so (e_n) is a monotone basic sequence in J by Corollary 4.1.25 and Proposition 4.1.21. Now fix a member (β_n) of J. To prove that (e_n) is a basis for J, it is enough to show that $\|(\beta_n)-\sum_{j=1}^k \beta_j e_j\|_a \to 0$ as $k \to \infty$, that is, that $\|(0,\dots,0,\beta_{k+1},\beta_{k+2},\dots)\|_a \to 0$ as $k \to \infty$. Suppose to the contrary that this did not happen. Then there would be a positive ϵ and, for each positive integer j, a collection $\{p_{1,j},\dots,p_{m(j),j}\}$ of positive integers such that $p_{1,j} < \cdots < p_{m(j),j} < p_{1,j+1}$ and

$$\sum_{n=1}^{m(j)-1} (\beta_{p_{n,j}} - \beta_{p_{n+1,j}})^2 + (\beta_{p_{m(j),j}} - \beta_{p_{1,j}})^2 \ge \epsilon.$$

Since $\beta_n \to 0$, there is some positive integer j_0 such that

$$\sum_{n=1}^{m(j)-1} (\beta_{p_{n,j}} - \beta_{p_{n+1,j}})^2 \ge \frac{\epsilon}{2}$$

when $j \ge j_0$. But a moment's thought about the definition of $\|\cdot\|_a$ then shows that $\|(\beta_n)\|_a = +\infty$, a contradiction. This finishes the proof that (e_n) is a basis for J and verifies the formula for basis expansions.

All that remains to be shown is that the monotone basis (e_n) is shrinking. Suppose it is not. Then there is an x^* in J^*, a positive δ, and a normalized block basic sequence (u_n) taken from (e_n) such that $x^*u_n \ge \delta$ for each n.

Therefore $\sum_n x^*(n^{-1}u_n)$ does not converge, so $\sum_n n^{-1}u_n$ does not converge. To obtain a contradiction to this and thereby finish the proof of the proposition, it is enough to show that there is a positive constant c such that $\|\sum_{n=m_1}^{m_2} n^{-1}u_n\|_b \leq c(\sum_{n=m_1}^{m_2} n^{-2})^{1/2}$ whenever $m_1 \leq m_2$.

Let (ζ_n) be a sequence of real numbers and (q_n) a sequence of positive integers such that $1 = q_1 < q_2 < q_3 < \cdots$ and $u_j = \sum_{n=q_j}^{q_{j+1}-1} \zeta_n e_n$ for each j. Notice that $|\zeta_n| \leq 1$ for each n since $\|u_j\|_a = 1$ for each j. A positive integer p will be said to *belong to* a term u_j of (u_n) if $q_j \leq p < q_{j+1}$.

Let positive integers m_1 and m_2 be such that $m_1 \leq m_2$, and let (γ_n) be the element $\sum_{n=m_1}^{m_2} n^{-1}u_n$ of J. Fix a positive integer m and positive integers $p_1, \ldots, p_m$ such that $m \geq 2$ and $p_1 < \cdots < p_m$, then let $s = \sum_{n=1}^{m-1}(\gamma_{p_n} - \gamma_{p_{n+1}})^2$. Let $\mathfrak{C}_1$ be the collection of summands $(\gamma_{p_k} - \gamma_{p_{k+1}})^2$ of s such that p_k and p_{k+1} both belong to the same term of (u_n), that is, such that there is some j for which $q_j \leq p_k < p_{k+1} < q_{j+1}$. For each member $(\gamma_{p_k} - \gamma_{p_{k+1}})^2$ of $\mathfrak{C}_1$ and the u_j to which it belongs, either

$$(\gamma_{p_k} - \gamma_{p_{k+1}})^2 = (j^{-1}\zeta_{p_k} - j^{-1}\zeta_{p_{k+1}})^2 = j^{-2}(\zeta_{p_k} - \zeta_{p_{k+1}})^2$$

or $(\gamma_{p_k} - \gamma_{p_{k+1}})^2 = 0$. It follows that if s_1 is the sum of all members of $\mathfrak{C}_1$, then

$$s_1 \leq 2\sum_{n=m_1}^{m_2} n^{-2}\|u_n\|_b^2 = 2\sum_{n=m_1}^{m_2} n^{-2}.$$

Now let $\mathfrak{C}_2$ be the collection of summands $(\gamma_{p_k} - \gamma_{p_{k+1}})^2$ of s such that p_k and p_{k+1} do not belong to the same term of (u_n), and let s_2 be the sum of all members of $\mathfrak{C}_2$. Fix a member $(\gamma_{p_k} - \gamma_{p_{k+1}})^2$ of $\mathfrak{C}_2$ and let u_{j_1} and u_{j_2} be the terms of (u_n) to which p_k and p_{k+1}, respectively, belong. Then γ_{p_k} is either $j_1^{-1}\zeta_{p_k}$ or 0, and similarly $\gamma_{p_{k+1}}$ is either $j_2^{-1}\zeta_{p_{k+1}}$ or 0. It follows that

$$\begin{aligned}(\gamma_{p_k} - \gamma_{p_{k+1}})^2 &\leq (j_1^{-1}\zeta_{p_k})^2 + (j_2^{-1}\zeta_{p_{k+1}})^2 + 2j_1^{-1}j_2^{-1}|\zeta_{p_k}\zeta_{p_{k+1}}| \\ &\leq j_1^{-2} + j_2^{-2} + 2j_1^{-1}j_2^{-1} \\ &\leq j_1^{-2} + j_2^{-2} + j_1^{-2} + j_2^{-2} \\ &= 2j_1^{-2} + 2j_2^{-2}.\end{aligned}$$

Of course, if $j_1 < m_1$ or $j_1 > m_2$, then j_1^{-1} and j_1^{-2} can be replaced by 0 in the above inequality, and similarly for j_2. Also, it is clear from the definition of $\mathfrak{C}_2$ that if the preceding argument is repeated for each member $(\gamma_{p_k} - \gamma_{p_{k+1}})^2$ of $\mathfrak{C}_2$, then each integer j_0 such that $m_1 \leq j_0 \leq m_2$ can appear in at most one of those arguments as j_1 and in at most one other as j_2. Consequently,

$$s_2 \leq 4\sum_{n=m_1}^{m_2} n^{-2}.$$

Therefore

$$2^{-1/2}\left(\sum_{n=1}^{m-1}(\gamma_{p_n}-\gamma_{p_{n+1}})^2\right)^{1/2}=\left(\frac{s_1+s_2}{2}\right)^{1/2}\leq 3^{1/2}\left(\sum_{n=m_1}^{m_2}n^{-2}\right)^{1/2}.$$

It follows from the definition of $\|\cdot\|_b$ that

$$\left\|\sum_{n=m_1}^{m_2}n^{-1}u_n\right\|_b\leq 3^{1/2}\left(\sum_{n=m_1}^{m_2}n^{-2}\right)^{1/2}.$$

As has already been shown, this is a contradiction that implies that (e_n) is shrinking. ■

As one would expect, the sequence (e_n) of standard unit vectors in J is called the *standard unit vector basis* for J.

The rest of this section is devoted to a close look at J^{**}. As observed in the comments following Proposition 4.5.2, the space J can be identified with the space of all sequences (α_n) of real numbers such that $(\sum_{n=1}^{m}\alpha_n e_n)_{m=1}^{\infty}$ converges, with the norm given by the formula

$$\|(\alpha_n)\|=\lim_m\left\|\sum_{n=1}^{m}\alpha_n e_n\right\|_a=\sup_m\left\|\sum_{n=1}^{m}\alpha_n e_n\right\|_a. \tag{4.2}$$

In fact, the space J is already precisely that space; in particular, the norm given by (4.2) is precisely $\|\cdot\|_a$. Furthermore, the space J^{**} can be viewed as the space of all sequences (α_n) of scalars such that $(\sum_{n=1}^{m}\alpha_n e_n)_{m=1}^{\infty}$ is bounded in J, with the norm also given by (4.2). For consistency, this norm will be denoted by $\|\cdot\|_a$ as for J. As an aid to understanding, in most of the rest of this section the space J^{**} will be treated as if it actually were this sequence space rather than just isometrically isomorphic to it, and J will often be viewed as the subspace of J^{**} with which the natural map from J into J^{**} identifies it; again, see the comments following Proposition 4.5.2. The reader wishing to rewrite the following arguments to include all of the appropriate isometric isomorphisms will have no trouble doing so.

Suppose that (β_n) is a sequence of reals such that $\lim_n\beta_n$ does not exist. Then there must be a positive ϵ and sequences (p_n) and (q_n) of positive integers such that $1\leq p_1<q_1<p_2<q_2<\cdots$ and such that $|\beta_{p_n}-\beta_{q_n}|\geq\epsilon$ for each n. For every positive integer m,

$$\left\|\sum_{n=1}^{q_m}\beta_n e_n\right\|_a\geq 2^{-1/2}\big((\beta_{p_1}-\beta_{q_1})^2+(\beta_{q_1}-\beta_{p_2})^2+\cdots$$
$$+(\beta_{p_m}-\beta_{q_m})^2+(\beta_{q_m}-\beta_{p_1})^2\big)^{1/2}\geq\left(\frac{m}{2}\right)^{1/2}\epsilon,$$

so $(\beta_n)\notin J^{**}$. Therefore $\lim_n\alpha_n$ exists whenever $(\alpha_n)\in J^{**}$.

Though $\|\cdot\|_a$ is being used to denote the norm of J^{**}, that does not mean that the formula of Definition 4.5.3 can be used to compute it. For example, it is easy to see that every constant sequence $(c, c, c, \dots)$ of real numbers is in J^{**} and has norm $|c|$ by (4.2), but the formula from Definition 4.5.3 would incorrectly give 0 for the value of the norm. For the moment, let the "norm" of a member (α_n) of J^{**} computed using the formula in Definition 4.5.3 be denoted by $\big|(\alpha_n)\big|_a$, and let $\|(\alpha_n)\|_a$ continue to represent the actual value of the norm of (α_n). For each (α_n) in J^{**},

$$2^{-1/2}\bigg(\sum_{n=1}^{m-1}(\alpha_{p_n} - \alpha_{p_{n+1}})^2 + (\alpha_{p_m} - \alpha_{p_1})^2\bigg)^{1/2} \le \bigg\|\sum_{n=1}^{p_m} \alpha_n e_n\bigg\|_a \le \|(\alpha_n)\|_a$$

whenever $m, p_1, \dots, p_m \in \mathbb{N}$, $m \ge 2$, and $p_1 < \cdots < p_m$, from which it follows that $\big|(\alpha_n)\big|_a \le \|(\alpha_n)\|_a$ and in particular that $\big|(\alpha_n)\big|_a$ is finite. One immediate consequence of this is that

$$J = \{\, (\alpha_n) : (\alpha_n) \in J^{**}, \lim_n \alpha_n = 0 \,\},$$

since the only way that a member of J^{**} can avoid being in J is to have a nonzero limit.

Incidentally, the formula for $\|\cdot\|_a$ of Definition 4.5.3 can be modified a bit so that the resulting slightly more complicated formula does also work to compute the norms of members of J^{**}. See Lemma 4.5.8.

Let $e_0 = (1, 1, 1, \dots)$. Then $e_0 \in J^{**}$ and $\|e_0\|_a = 1$. It will now be shown that $(e_0, e_1, e_2, \dots)$ is a basis for J^{**}. Fix an element (γ_n) of J^{**} and let $\gamma = \lim_n \gamma_n$. Then $\lim_n(\gamma_n - \gamma) = 0$, so $(\gamma_n) - \gamma e_0 \in J$ and $(\gamma_n) - \gamma e_0 = \sum_{n=1}^{\infty}(\gamma_n - \gamma)e_n$. Therefore

$$(\gamma_n) = \gamma e_0 + \sum_{n=1}^{\infty}(\gamma_n - \gamma)e_n.$$

Now suppose that $(\zeta_n)_{n=0}^{\infty}$ is any sequence of reals such that

$$(\gamma_n) = \zeta_0 e_0 + \sum_{n=1}^{\infty} \zeta_n e_n.$$

Then $(\gamma_n) - \zeta_0 e_0 = \sum_{n=1}^{\infty} \zeta_n e_n \in J$, which implies that $\lim_n(\gamma_n - \zeta_0) = 0$, which in turn implies that $\zeta_0 = \gamma$. Therefore

$$\sum_{n=1}^{\infty} \zeta_n e_n = \sum_{n=1}^{\infty}(\gamma_n - \gamma)e_n \in J,$$

so $\zeta_n = \gamma_n - \gamma$ when $n \ge 1$. Thus, there is one and only one sequence $(\zeta_n)_{n=0}^{\infty}$ of reals such that $(\gamma_n) = \sum_{n=0}^{\infty} \zeta_n e_n$, and so $(e_n)_{n=0}^{\infty}$ is a basis for J^{**}.

It would not hurt to pause for a moment to recap in rigorous language part of what has been discovered in the preceding discussion.

4.5.6 Proposition. *Let $X_{J^{**}}$ be the Banach space of sequences of scalars described above that is identified with J^{**} and has J as a subspace, and let $\Lambda\colon X_{J^{**}} \to J^{**}$ be the identifying isometric isomorphism. Let (e_n) be the standard unit vector basis for J, and let e_0 be the element $(1,1,1,\dots)$ of $X_{J^{**}}$. Let Q be the natural map from J into J^{**}.*

(a) *If $(\alpha_n) \in J$, then $Q(\alpha_n) = \Lambda(\alpha_n)$.*

(b) *If $(\alpha_n) \in X_{J^{**}}$, then $\lim_n \alpha_n$ exists.*

(c) $J = \{\,(\alpha_n) : (\alpha_n) \in X_{J^{**}},\ \lim_n \alpha_n = 0\,\}$.

(d) *The sequence $(e_n)_{n=0}^{\infty}$ is a basis for $X_{J^{**}}$. Consequently, the sequence $(\Lambda e_n)_{n=0}^{\infty}$, which is the same as $(\Lambda e_0, Qe_1, Qe_2, \dots)$, is a basis for J^{**}.*

Recall that if W is a vector subspace of a vector space V, then the *codimension* of W in V is the dimension of the quotient vector space V/W; that is, roughly speaking, the number of dimensions that W lacks of being all of V. Let J^{**} be identified as before with a space of sequences of scalars and J with its natural image in J^{**}. Suppose that $x^{**} \in J^{**}$ and that $\sum_{n=0}^{\infty} \zeta_n e_n$ is the expansion for x^{**} in terms of the basis $(e_n)_{n=0}^{\infty}$ for J^{**}. Consider the element $x^{**} + J$ of the quotient Banach space J^{**}/J. Since $\sum_{n=1}^{\infty} \zeta_n e_n \in J$, it follows that $x^{**} + J = \zeta_0 e_0 + J$, so $J^{**}/J = \langle\{e_0 + J\}\rangle$. Therefore $\{e_0 + J\}$ is a basis for J^{**}/J, so J has codimension 1 in J^{**}. This can be stated more formally as follows.

4.5.7 Theorem. *Let Q be the natural map from J into J^{**}. Then the codimension of $Q(J)$ in J^{**} is 1. Thus, the space J is not reflexive.*

Though the natural map is an isometric isomorphism from J onto a proper subspace of J^{**}, there is another isometric isomorphism from J onto all of J^{**}. The key to the proof of that to be given here is the following formula for computing the norms of members of J^{**}.

4.5.8 Lemma. *Let J^{**} be identified with a Banach space of sequences of scalars as has been done above, and let (α_n) be a member of J^{**}. Then $\|(\alpha_n)\|_a$ is the supremum over all real numbers t such that t equals either*

$$2^{-1/2}\Big(\sum_{n=1}^{m-1}(\alpha_{p_n} - \alpha_{p_{n+1}})^2 + (\alpha_{p_m} - \alpha_{p_1})^2\Big)^{1/2} \tag{4.3}$$

or

$$2^{-1/2}\Big(\sum_{n=1}^{m-2}(\alpha_{p_n} - \alpha_{p_{n+1}})^2 + \alpha_{p_{m-1}}^2 + \alpha_{p_1}^2\Big)^{1/2} \tag{4.4}$$

for some positive integers $m, p_1, \dots, p_m$ such that $m \ge 2$ and $p_1 < \cdots < p_m$. (By convention, the sum from 1 to $m-2$ is considered to be 0 if $m = 2$.)

PROOF. For each positive integer k, let $\mathbf{I}_{\{1,\ldots,k\}}$ be the indicator function of the set $\{1,\ldots,k\}$. Then

$$\begin{aligned}
\|(\alpha_n)\|_a &= \sup_k \Big\| \sum_{n=1}^{k} \alpha_n e_n \Big\|_a \\
&= \sup_k \|(\alpha_1,\ldots,\alpha_k,0,0,\ldots)\|_a \\
&= \sup\Big\{ 2^{-1/2}\Big(\sum_{n=1}^{m-2} (\alpha_{p_n} - \alpha_{p_{n+1}})^2 + \big(\alpha_{p_{m-1}} - \alpha_{p_m}\mathbf{I}_{\{1,\ldots,k\}}(p_m)\big)^2 \\
&\qquad\qquad + \big(\alpha_{p_m}\mathbf{I}_{\{1,\ldots,k\}}(p_m) - \alpha_{p_1}\big)^2 \Big)^{1/2} : \\
&\qquad\qquad\qquad m \geq 2,\ k \in \mathbb{N},\ p_1 < \cdots < p_m,\ p_{m-1} \leq k \Big\}.
\end{aligned}$$

A moment's thought shows that this last supremum is precisely the one described in the statement of this lemma. ■

For the duration of this paragraph, let (α_n) be a fixed member of J^{**}. It is easy to see that $\|(0,\alpha_1,\ldots,\alpha_m,0,0,\ldots)\|_a = \|(\alpha_1,\ldots,\alpha_m,0,0,\ldots)\|_a$ for each m, from which it follows that

$$\sup_m \Big\| \sum_{n=1}^{m} \alpha_n e_{n+1} \Big\|_a = \sup_m \Big\| \sum_{n=1}^{m} \alpha_n e_n \Big\|_a = \|(\alpha_n)\|_a < +\infty$$

and therefore that the "shifted" sequence $(0,\alpha_1,\alpha_2,\ldots)$ is in J^{**}. Subtracting $(\lim_n \alpha_n)e_0$ from this shifted sequence yields the member

$$(-\lim_n \alpha_n,\ \alpha_1 - \lim_n \alpha_n,\ \alpha_2 - \lim_n \alpha_n,\ \ldots)$$

of J^{**}, which is in fact in J since the limit of this sequence is 0. If the norm of this member of J is computed using the formula of Definition 4.5.3 (with the $2^{-1/2}$ brought inside the supremum), then the set whose supremum is being taken is precisely the set whose supremum is used to compute $\|(\alpha_n)\|_a$ in Lemma 4.5.8, so the above member of J has the same norm as (α_n).

Define $T\colon J^{**} \to J$ by the formula

$$T\big((\alpha_n)\big) = (-\lim_n \alpha_n,\ \alpha_1 - \lim_n \alpha_n,\ \alpha_2 - \lim_n \alpha_n,\ \ldots).$$

It is clear that this map is linear, and therefore is an isometric isomorphism from J^{**} into J by the result of the preceding paragraph. Now suppose that $(\beta_n) \in J$. Then $(\beta_2,\beta_3,\ldots)$ is clearly in J, so $(\beta_2,\beta_3,\ldots) - \beta_1 e_0 \in J^{**}$. It is easy to check that $T\big((\beta_2,\beta_3,\ldots) - \beta_1 e_0\big) = (\beta_n)$. Therefore $T(J^{**}) = J$, which finishes the proof of the following result.

4.5.9 Theorem. *The Banach space J is isometrically isomorphic to J^{**}.*

The ideas and results of this section are essentially from R. C. James's papers [109] and [110] that appeared in 1950 and 1951 respectively. In the earlier paper, James constructed a Banach space that is almost J and is isomorphic to it; however, the norm is not quite right for the space to be *isometrically* isomorphic to its second dual, so James had to settle for an isomorphism. By tweaking the norm a bit in the later paper, he was able to obtain the stronger result.

The space J proved useful as a counterexample for several long-standing conjectures. As was discussed at the end of Section 1.11, a comment in Banach's book [13] can be interpreted to be the question of whether a separable Banach space must be reflexive if there is any isometric isomorphism whatever from the space onto its second dual. The space J shows that the answer is no. Also, the separability of J^{**} shows that J is a counterexample for the conjecture that a Banach space must be reflexive if its second dual is separable.[4] See Exercises 4.44 and 4.45 for two other conjectures settled in the negative by J. A list of references for other applications of J can be found in [118, p. 633].

Exercises

4.43 Prove that J has no unconditional basis.

4.44 Prove that J is an example of an infinite-dimensional Banach space X such that X is not isomorphic to $X \oplus X$. Exercise 3.12 could help.

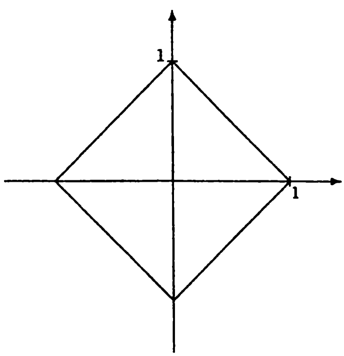
4.45 Prove that J is an infinite-dimensional real Banach space for which there is no complex Banach space X such that J is isomorphic to the real Banach space X_r obtained from X by restricting multiplication of vectors by scalars to $\mathbb{R} \times X$. It might help to look at Exercise 3.12 as well as Proposition 1.13.1 and its proof.

[4]However, it was not quite the first such counterexample. That honor went to J's isomorphic cousin constructed in [109], as James observed in that paper.

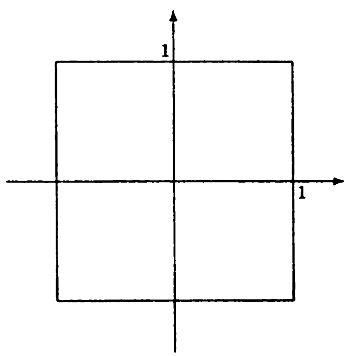
5

Rotundity and Smoothness

When thinking of the closed unit ball of a normed space, it is tempting to visualize some round, smooth object like the closed unit ball of real Euclidean 2- or 3-space. However, closed unit balls are sometimes not so nicely shaped. Consider, for example, the closed unit balls of real ℓ_1^2 and ℓ_∞^2. Neither is round by any of the usual meanings of that word, since their boundaries,



$S_{\ell_1^2}$



$S_{\ell_\infty^2}$

which is to say the unit spheres of the spaces, are each composed of four straight line segments. Also, neither is smooth along its entire boundary, since each has four corners. These features of the closed unit balls have a number of interesting consequences that cause the norms of these two spaces to behave a bit unlike that of real Euclidean 2-space. For example, if z_1 and z_2 are different points on any one of the four sides of one of these balls, then $1 = \|z_1\| = \|z_2\| = \frac{1}{2}\|z_1\| + \frac{1}{2}\|z_2\| = \left\|\frac{1}{2}z_1 + \frac{1}{2}z_2\right\|$, so equality is attained in the inequality $\|z_1 + z_2\| \leq \|z_1\| + \|z_2\|$ despite the fact that neither z_1 nor z_2 is a nonnegative real multiple of the other. Furthermore, the presence of the corners leads to the existence of multiple *norming functionals* for some points z of the unit sphere of each of these spaces, that is, norm-one members z^* of the dual space such that $z^*z = \|z\|$. To see why this would be, let Z be either of these two spaces and let z_0 be one of the four corners of B_Z. Then there are infinitely many different straight lines that pass through z_0 without intersecting the interior of the closed unit ball; let l_1 and l_2 be two of them. By Mazur's separation theorem, there are members z_1^* and z_2^* of Z^*, necessarily different, such that if $j \in \{1, 2\}$, then $z_j^*z = 1$ when $z \in l_j$ and $z_j^*z \leq 1$ when $z \in B_Z$. It follows readily that z_1^* and z_2^* are both norming functionals for z_0. As will be seen, it is precisely the presence of the corners or sharp bends in the unit sphere that caused this multiplicity of norming functionals for elements at the locations of the bends.

The purpose of this chapter is to study the special properties of normed spaces whose closed unit balls are round, in the sense that the unit spheres include no nontrivial line segments, and of those whose closed unit balls are smooth, in the sense that the unit spheres have no corners or sharp bends. Roundness will be taken up first.

5.1 Rotundity

The property about to be defined was formulated independently by James Clarkson and Mark Krein. Clarkson was particularly interested in the uniform version of this property that will be studied in the next section, while Krein used it in joint work with Naum Akhiezer on the moment problem. There is an entire host of equivalent ways to define this property, some of which are given in this section.

5.1.1 Definition. (J. A. Clarkson, 1936 [41]; N. I. Akhiezer and M. G. Krein, 1938 [1]). A normed space X is *rotund* or *strictly convex* or *strictly normed* if $\|tx_1 + (1-t)x_2\| < 1$ whenever x_1 and x_2 are different points of S_X and $0 < t < 1$.

If a normed space satisfies the condition of the preceding definition, then it is sometimes said that its norm or its closed unit ball is rotund rather than attaching that label to the entire space. The term *rotund* used here is due to Mahlon Day, who published a series of six important papers

[47]–[52] on rotundity and uniform rotundity between 1941 and 1957.[1] The term *strictly normed* is Krein's.

Definition 5.1.1 actually says that a normed space is rotund when its unit sphere includes no nontrivial straight line segments, but some arguing needs to be done to show that. Suppose that X is a normed space. Just for the remainder of this paragraph, let $(x_1; x_2)$ denote the "open line segment" $\{ tx_1 + (1-t)x_2 : 0 < t < 1 \}$ whenever $x_1, x_2 \in X$. If X is rotund and x_1 and x_2 are different points of its unit sphere, then the definition of rotundity assures that $(x_1; x_2)$ lies entirely in the interior of B_X, so no nontrivial line segments lie in S_X. Suppose conversely that no nontrivial line segments lie in S_X and that x_1 and x_2 are different points of S_X. Then some of the points of $(x_1; x_2)$ lie in B_X°, so it follows easily from Lemma 2.2.18 that *all* of the points of $(x_1; x_2)$ lie in B_X°. The space X is therefore rotund. Incidentally, the last portion of this argument also establishes the following characterization of rotundity that allows the verification of the property by examining only the midpoints of straight line segments rather than the entire segments.

5.1.2 Proposition. *Suppose that X is a normed space. Then X is rotund if and only if $\left\|\frac{1}{2}(x_1 + x_2)\right\| < 1$ whenever x_1 and x_2 are different points of S_X.*

It is time for some examples.

5.1.3 Example. The scalar field $\mathbb{F}$, viewed as a normed space over $\mathbb{F}$, is obviously rotund. More generally, it is easy to see that every normed space that is zero- or one-dimensional is rotund.

5.1.4 Example. Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω and that $1 < p < \infty$. Then $L_p(\Omega, \Sigma, \mu)$ is rotund. The argument showing this will be saved for the proof of Theorem 5.2.11, in which it will be shown that $L_p(\Omega, \Sigma, \mu)$ has the stronger property of uniform rotundity.

5.1.5 Example. This generalizes the examples of nonrotund spaces given in the introduction to this chapter. As in the last example, suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω, but suppose in addition that there are two disjoint measurable subsets A_1 and A_2 of Ω each having finite positive measure. Let $\mathbf{I}_{A_1}$ and $\mathbf{I}_{A_2}$ be the respective indicator functions of these sets, and let $f_1 = \left(\mu(A_1)\right)^{-1}\mathbf{I}_{A_1}$, $f_2 = \left(\mu(A_2)\right)^{-1}\mathbf{I}_{A_2}$,

---

[1] However, Day referred to the property as strict convexity in those papers, as most others did before and, admittedly, have since. The term *rotund* first appeared in his 1958 book [53].

$g_1 = \mathbf{I}_{A_1} + \mathbf{I}_{A_2}$, and $g_2 = \mathbf{I}_{A_1} - \mathbf{I}_{A_2}$. It is easy to check that

$$\|f_1\|_1 = \|f_2\|_1 = \left\|\tfrac{1}{2}(f_1+f_2)\right\|_1 = \|g_1\|_\infty = \|g_2\|_\infty = \left\|\tfrac{1}{2}(g_1+g_2)\right\|_\infty = 1,$$

from which it follows that neither $L_1(\Omega,\Sigma,\mu)$ nor $L_\infty(\Omega,\Sigma,\mu)$ is rotund. In particular, the spaces ℓ_1 and ℓ_∞ are not rotund, and ℓ_1^n and ℓ_∞^n are not rotund when $n \geq 2$. Notice that ℓ_1^n and ℓ_∞^n are rotund when $n < 2$ by Example 5.1.3.

5.1.6 Example. Let e_1 and e_2 be the first two standard unit vectors of c_0. Let $x_1 = e_1 + e_2$ and $x_2 = e_1 - e_2$. Then

$$\|x_1\|_\infty = \|x_2\|_\infty = \left\|\tfrac{1}{2}(x_1+x_2)\right\|_\infty = 1,$$

so neither c_0 nor ℓ_∞ is rotund.

5.1.7 Example. Suppose that K is a compact Hausdorff space having more than one element. Let k_1 and k_2 be different members of K. Then Urysohn's lemma assures that there is a continuous function $f_1 : K \to [0,1]$ such that $f_1(k_1) = 0$ and $f_1(k_2) = 1$. Let f_2 be the function taking on the value 1 everywhere on K. Then

$$\|f_1\|_\infty = \|f_2\|_\infty = \left\|\tfrac{1}{2}(f_1+f_2)\right\|_\infty = 1,$$

so $C(K)$ is not rotund.

5.1.8 Example. The evidence presented so far might lead one to conjecture that rotund Banach spaces are always reflexive. The purpose of this example is to construct a nonreflexive rotund Banach space by finding a rotund norm for ℓ_1 that is equivalent to its original norm. This construction might seem to have some unnecessary complications in it, but the new norm is being designed to have some special properties that will be useful in Example 5.1.22.

The first step is to define some functions. For each positive integer m and each nonnegative real number t, let

$$f_m(t) = \frac{t^2 + mt}{m+1}.$$

Then for each m,

(1) the function f_m maps $[0,1]$ continuously onto $[0,1]$, and $f_m(0) = 0$ and $f_m(1) = 1$;

(2) f_m, f_m', and f_m'' are all positive on $(0,\infty)$, so in particular the function f_m is strictly increasing on $[0,\infty)$;

(3) $\frac{m}{m+1}t \leq f_m(t) \leq t$ whenever $t \in [0,1]$, with strict inequality when $t \in (0,1)$;

(4) $f_m(st_1+(1-s)t_2) < sf_m(t_1)+(1-s)f_m(t_2)$ whenever t_1 and t_2 are different members of $[0,\infty)$ and $0<s<1$; and

(5) $|f_m(t_1)-f_m(t_2)| \le \max\{f_m'(t) : 0\le t\le 1\}\,|t_1-t_2| \le \frac{3}{2}|t_1-t_2|$ whenever $t_1,t_2\in[0,1]$.

Let $\{A_m : m\in\mathbb{N}\}$ be a collection of pairwise disjoint infinite subsets of $\mathbb{N}$ whose union is $\mathbb{N}$, and for each positive integer n let $m(n)$ be the index m of the set A_m containing n. Let

$$C=\Big\{(\alpha_n) : (\alpha_n)\in\ell_1,\ \sum_n f_{m(n)}(|\alpha_n|)\le 1\Big\}.$$

Notice that if $(\alpha_n)\in C$, then $|\alpha_n|\le 1$ for each n. Suppose that $((\zeta_{n,j}))_{j=1}^{\infty}$ is a sequence in C that converges to some member (ζ_n) of ℓ_1. For each j,

$$\begin{aligned}\Big|\sum_n f_{m(n)}(|\zeta_{n,j}|)-\sum_n f_{m(n)}(|\zeta_n|)\Big| &\le \sum_n \big|f_{m(n)}(|\zeta_{n,j}|)-f_{m(n)}(|\zeta_n|)\big|\\ &\le \tfrac{3}{2}\sum_n \big|\,|\zeta_{n,j}|-|\zeta_n|\,\big|\\ &\le \tfrac{3}{2}\sum_n |\zeta_{n,j}-\zeta_n|\\ &= \tfrac{3}{2}\|(\zeta_{n,j})-(\zeta_n)\|_1,\end{aligned}$$

so $\sum_n f_{m(n)}(|\zeta_n|)=\lim_j\sum_n f_{m(n)}(|\zeta_{n,j}|)\le 1$ and $(\zeta_n)\in C$. The set C is therefore closed. If $(\beta_n),(\gamma_n)\in C$ and $0<s<1$, then

$$\begin{aligned}\sum_n f_{m(n)}(|s\beta_n+(1-s)\gamma_n|) &\le \sum_n f_{m(n)}(s|\beta_n|+(1-s)|\gamma_n|)\\ &\le s\sum_n f_{m(n)}(|\beta_n|)+(1-s)\sum_n f_{m(n)}(|\gamma_n|)\\ &\le 1,\end{aligned}$$

so $s(\beta_n)+(1-s)(\gamma_n)\in C$. The set C is therefore convex. It is clear that C is balanced and, since $B_{\ell_1}\subseteq C$, absorbing. The Minkowski functional p_C of C is therefore a seminorm on ℓ_1 by Proposition 1.9.14. Now $f_m(t)\ge t/2$ whenever $m\in\mathbb{N}$ and $0\le t\le 1$. From this and Proposition 1.9.14, it follows that

$$\{x : x\in\ell_1,\ p_C(x)<1\}\subseteq C\subseteq 2B_{\ell_1},$$

which in turn implies that $p_C(x)=0$ only if $x=0$. The seminorm p_C is therefore actually a norm, which will be denoted by $\|\cdot\|_r$. The normed space resulting from this new norm will be denoted by $\ell_{1,r}$ to distinguish it from $(\ell_1,\|\cdot\|_1)$. Since $B_{\ell_1}\subseteq C\subseteq 2B_{\ell_1}$, it is easy to check that

$\frac{1}{2}\|x\|_1 \le \|x\|_r \le \|x\|_1$ whenever $x \in \ell_1$, so the norms $\|\cdot\|_1$ and $\|\cdot\|_r$ are equivalent. Now

$$\{x : x \in \ell_1, \|x\|_r < 1\} \subseteq C \subseteq \{x : x \in \ell_1, \|x\|_r \le 1\}$$

by Proposition 1.9.14, which together with the fact that C is closed in the common topology of ℓ_1 and $\ell_{1,r}$ shows that C is the closed unit ball of $\ell_{1,r}$.

The equivalence of $\|\cdot\|_1$ and $\|\cdot\|_r$ shows that $\ell_{1,r}$ is a nonreflexive Banach space, so all that remains to be proved is that the norm $\|\cdot\|_r$ is rotund. Suppose that (β_n) and (γ_n) are different elements of $S_{\ell_{1,r}}$. Let $(\mu_n) = \frac{1}{2}\big((\beta_n) + (\gamma_n)\big)$. It is enough to find a real number a greater than 1 such that $a(\mu_n) \in C$, for this will imply that $\|a(\mu_n)\|_r \le 1$ and therefore that $\|(\mu_n)\|_r < 1$. To this end, let n_0 be a positive integer such that $\beta_{n_0} \ne \gamma_{n_0}$. Then either $\left|\frac{1}{2}(\beta_{n_0} + \gamma_{n_0})\right| < \frac{1}{2}|\beta_{n_0}| + \frac{1}{2}|\gamma_{n_0}|$ or $|\beta_{n_0}| \ne |\gamma_{n_0}|$. It follows that one of the first two inequality symbols in the inequality

$$\begin{aligned}\sum_n f_{m(n)}\big(|\mu_n|\big) &= \sum_n f_{m(n)}\Big(\big|\tfrac{1}{2}(\beta_n + \gamma_n)\big|\Big)\\ &\le \sum_n f_{m(n)}\big(\tfrac{1}{2}|\beta_n| + \tfrac{1}{2}|\gamma_n|\big)\\ &\le \tfrac{1}{2}\sum_n f_{m(n)}\big(|\beta_n|\big) + \tfrac{1}{2}\sum_n f_{m(n)}\big(|\gamma_n|\big)\\ &\le 1\end{aligned}$$

actually represents a strict inequality, so $\sum_n f_{m(n)}\big(|\mu_n|\big) < 1$. This implies that $\max_n|\mu_n| < 1$, so there is a real number a greater than 1 such that $\max_n|a\mu_n| < 1$. An argument like that used to demonstrate that C is closed then shows that

$$\begin{aligned}\left|\sum_n f_{m(n)}\big(|a\mu_n|\big) - \sum_n f_{m(n)}\big(|\mu_n|\big)\right| &\le \tfrac{3}{2}\|a(\mu_n) - (\mu_n)\|_1\\ &= \tfrac{3}{2}(a-1)\|(\mu_n)\|_1,\end{aligned}$$

so it is possible to reduce a enough that $\sum_n f_{m(n)}\big(|a\mu_n|\big) < 1$ while leaving a greater than 1. It follows that $a(\mu_n) \in C$, which establishes the rotundity of $\ell_{1,r}$.

It should be noted that in the last few lines of the preceding argument it was established that if $(\alpha_n) \in \ell_{1,r}$ and $\sum_n f_{m(n)}\big(|\alpha_n|\big) < 1$, then $\|(\alpha_n)\|_r < 1$. This fact will be used again in Example 5.1.22.

One consequence of Examples 5.1.4 and 5.1.5 (which, for the real case at least, is also apparent from diagrams of the corresponding unit spheres) is that Euclidean 2-space is rotund while ℓ_1^2 is not, even though the two spaces are isomorphic. Thus, in contrast to most properties of normed spaces that have been studied so far, rotundity is not isomorphism-invariant. The following is the most that can be said along these lines. The proof is obvious.

5.1.9 Proposition. *Every normed space that is isometrically isomorphic to a rotund normed space is itself rotund.*

The next portion of this section is devoted to obtaining a number of characterizations of rotundity. One easy one involves extreme points and is based on the observation that if x_0 is a point in the unit sphere of a normed space X and there are elements x_1 and x_2 of B_X and a real number t such that $0 < t < 1$ and $x_0 = tx_1 + (1-t)x_2$, then x_1 and x_2 also lie in S_X; this is easy to prove directly and also follows immediately from Lemma 2.2.18. Thus, if a point in S_X is an "interior point" of a "closed line segment" $\{\, tx_1 + (1-t)x_2 : 0 \le t \le 1 \,\}$ in B_X such that $x_1 \ne x_2$, then both endpoints lie in S_X, which by Lemma 2.2.18 implies that the entire "closed line segment" lies in S_X. Therefore every point of S_X is an extreme point of B_X if and only if S_X includes no nontrivial "closed line segments," which happens if and only if X is rotund.

5.1.10 Proposition. *A normed space is rotund if and only if each element of its unit sphere is an extreme point of its closed unit ball.*

Another useful characterization of rotundity is based on the occurrence of equality in the triangle inequality. It is an elementary fact from Euclidean geometry that if A, B, and C are points in the Euclidean plane and B does not lie on the straight line segment connecting A to C, then the distance from A to C must be strictly less than the sum of the distances from A to B and from B to C. Treating the vectors of real Euclidean 2-space as arrows[2] transforms this property of the Euclidean plane into the statement that if $x_1, x_2 \in \ell_2^2$, then equality occurs in the inequality $\|x_1 + x_2\|_2 \le \|x_1\|_2 + \|x_2\|_2$ only if one of the vectors x_1 and x_2 is a nonnegative real multiple of the other. It turns out that this condition characterizes rotund norms among all norms.

5.1.11 Proposition. *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is rotund.*

(b) *Whenever $x_1, x_2 \in X$ and $\|x_1 + x_2\| = \|x_1\| + \|x_2\|$, one of the two vectors must be a nonnegative real multiple of the other.*

PROOF. Suppose first that X is rotund and that x_1 and x_2 are members of X such that $\|x_1 + x_2\| = \|x_1\| + \|x_2\|$. It must be shown that one of the two vectors x_1 and x_2 is a nonnegative multiple of the other, so it may be assumed that neither is zero. It may even be assumed that $1 = \|x_1\| \le \|x_2\|$.

---

[2]See the introduction to Section 1.2.

Let $y = \|x_2\|^{-1}x_2$. Then

$$\begin{aligned} 2 &\geq \|x_1 + y\| \\ &= \left\|x_1 + x_2 - \left(1 - \|x_2\|^{-1}\right)x_2\right\| \\ &\geq \|x_1 + x_2\| - \left(1 - \|x_2\|^{-1}\right)\|x_2\| \\ &= \|x_1\| + \|x_2\| - \|x_2\| + 1 \\ &= 2, \end{aligned}$$

from which it follows that $\left\|\frac{1}{2}(x_1 + y)\right\| = 1$. Since $x_1, y \in S_X$, the rotundity of X requires that $x_1 = y = \|x_2\|^{-1}x_2$. This shows that (a) $\Rightarrow$ (b).

Suppose conversely that (b) holds. Let z_1 and z_2 be different members of S_X. Then neither of the two vectors is a nonnegative multiple of the other, which implies that $\|z_1 + z_2\| < \|z_1\| + \|z_2\| = 2$ and therefore that $\left\|\frac{1}{2}(z_1 + z_2)\right\| < 1$. The space X is therefore rotund, so (b) $\Rightarrow$ (a). ∎

It should be noted that for any normed space whatever and any two members x_1 and x_2 of the space such that one of the vectors is a nonnegative multiple of the other, it is true that $\|x_1 + x_2\| = \|x_1\| + \|x_2\|$. Thus, the preceding proposition yields the following strengthened form of the triangle inequality for rotund normed spaces.

5.1.12 Corollary. *Suppose that X is a rotund normed space and that $x_1, x_2 \in X$. Then $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$, with equality if and only if one of the two vectors is a nonnegative real multiple of the other.*

Still another characterization of rotundity involves support hyperplanes. Most of the following definition was given earlier as Definition 2.11.3, but will be repeated here since it appeared in an optional section.

5.1.13 Definition. Let A be a subset of a topological vector space X. A nonzero x^* in X^* is a *support functional* for A if there is an x_0 in A such that $\operatorname{Re} x^*x_0 = \sup\{\operatorname{Re} x^*x : x \in A\}$, in which case x_0 is a *support point* of A, the set $\{x : x \in X, \operatorname{Re} x^*x = \operatorname{Re} x^*x_0\}$ is a *support hyperplane* for A, and the functional x^* and the support hyperplane are both said to *support* A at x_0.

Notice that if H is the support hyperplane of the preceding definition, then $H \cap A$ is exactly the collection of points at which H supports A, and this intersection is by definition nonempty.

An abuse of terminology is being committed in the preceding definition. In standard linear algebra parlance, a hyperplane in a vector space V is usually defined to be a set of the form $\{v : v \in V, f(v) = \alpha_0\}$, where f is a nonzero linear functional on V and α_0 is a scalar. In Definition 5.1.13, the support "hyperplane" is the set of all points in X where a particular *real*-linear functional, rather than a true linear functional, takes on some constant value. Of course, there is no difference when the scalar field is $\mathbb{R}$, but there is when it is $\mathbb{C}$. For example, let x^* be the linear functional on the complex Banach space $\mathbb{C}$ given by the formula $x^*\alpha = \alpha$. Then x^* supports $B_{\mathbb{C}}$ at 1, with the corresponding support hyperplane being the line $\{\alpha : \alpha \in \mathbb{C}, \operatorname{Re}\alpha = 1\}$. However, it is easy to see that the true hyperplanes in $\mathbb{C}$ are exactly the individual points of $\mathbb{C}$ rather than lines. Of course, every complex normed space X is also a real normed space when multiplication of vectors by scalars is restricted to $\mathbb{R} \times X$. Thus, when X is a complex normed space, the object called a hyperplane in Definition 5.1.13 really is a hyperplane in the standard linear algebra sense if X is treated as a real normed space in this fashion.

To see what motivates the next characterization of rotundity, it will be helpful to review some of the properties of the hyperplanes of the real vector space $\mathbb{R}^2$. Since every linear functional on $\mathbb{R}^2$ has the form $(\alpha, \beta) \mapsto \alpha_0\alpha + \beta_0\beta$ where $\alpha_0, \beta_0 \in \mathbb{R}$, the hyperplanes in $\mathbb{R}^2$ are exactly the sets of the form

$$\{(\alpha, \beta) : (\alpha, \beta) \in \mathbb{R}^2, \ \alpha_0\alpha + \beta_0\beta = \gamma_0\}$$

such that $\alpha_0, \beta_0, \gamma_0 \in \mathbb{R}$ and α_0 and β_0 are not both zero. Thus, the hyperplanes in $\mathbb{R}^2$ are precisely the straight lines in $\mathbb{R}^2$. The hyperplanes in $\mathbb{R}^2$ that pass through the origin, that is, the *hyperspaces* in $\mathbb{R}^2$, are precisely the kernels of the linear functionals on the space. If H is a hyperplane in $\mathbb{R}^2$ that does not pass through the origin, then clearly there is some linear functional f on $\mathbb{R}^2$ such that

$$H = \{(\alpha, \beta) : (\alpha, \beta) \in \mathbb{R}^2, \ f(\alpha, \beta) = 1\},$$

from which it follows that

$$\{(\alpha, \beta) : (\alpha, \beta) \in \mathbb{R}^2, \ f(\alpha, \beta) < 1\}$$

is the open halfspace bounded by H that contains the origin, while

$$\{(\alpha, \beta) : (\alpha, \beta) \in \mathbb{R}^2, \ f(\alpha, \beta) > 1\}$$

is the open halfspace on the other side of H.

Now suppose that X is a real normed space whose underlying vector space is $\mathbb{R}^2$. It follows from the preceding discussion that the support hyperplanes for B_X are exactly the straight lines in $\mathbb{R}^2$ that intersect B_X without penetrating its interior, that is, the "tangent lines" to B_X. If X is rotund, then no support hyperplane for B_X can intersect B_X at more than one point, for the straight line segment connecting two points of intersection would have to lie in S_X. Suppose conversely that X is not rotund. Let x_1 and x_2 be different points of S_X such that the entire line segment $\{tx_1 + (1-t)x_2 : 0 \le t \le 1\}$ lies in S_X. By Lemma 2.2.18, the straight line that includes this segment cannot intersect B_X°, so this line is a support hyperplane for B_X that intersects it at more than one point. To summarize, the space X is rotund if and only if each support hyperplane for B_X intersects it at exactly one point. This fact generalizes to arbitrary normed spaces.

5.1.14 Lemma. *Suppose that H is a support hyperplane for a subset A of a topological vector space. Then $H \cap A^\circ = \varnothing$.*

PROOF. Let X be the topological vector space in question, and let x^* be a nonzero member of X^* and x_0 an element of A such that $\operatorname{Re} x^*x_0 = \sup\{\operatorname{Re} x^*x : x \in A\}$ and $H = \{x : x \in X, \operatorname{Re} x^*x = \operatorname{Re} x^*x_0\}$. Suppose that there is some y in $H \cap A^\circ$. Let x_1 be an element of X such that $\operatorname{Re} x^*x_1 = 1$. Then $y + \delta x_1 \in A^\circ$ for some small positive δ, so

$$\operatorname{Re} x^*x_0 \geq \operatorname{Re} x^*(y + \delta x_1) = \operatorname{Re} x^*x_0 + \delta > \operatorname{Re} x^*x_0,$$

a contradiction. ■

5.1.15 Theorem. (A. F. Ruston, 1949 [203]). *A normed space X is rotund if and only if each support hyperplane for B_X supports B_X at only one point.*

PROOF. This proof is very much like the one done above for real normed spaces having $\mathbb{R}^2$ as their underlying vector space. Suppose first that X is rotund. Let H be a support hyperplane for B_X. Then $H \cap B_X \subseteq S_X$ since $H \cap B_X^\circ = \varnothing$. The convexity of H and B_X assures that $H \cap B_X$ is also convex, so $H \cap B_X$ cannot have two distinct points in it without also including the entire straight line segment connecting the points. However, the set S_X includes no nontrivial straight line segments, so the only point in $H \cap B_X$ is the one guaranteed by the definition of a support hyperplane.

Suppose conversely that X is not rotund. Let x_1 and x_2 be distinct elements of S_X such that $\{tx_1 + (1-t)x_2 : 0 \leq t \leq 1\} \subseteq S_X$; call this line segment C. By Eidelheit's separation theorem, there is an x^* in X^* such that $\operatorname{Re} x^*x \geq 1$ for each x in C and $\operatorname{Re} x^*x \leq 1$ for each x in B_X. Notice in particular that $\operatorname{Re} x^*x_1 = \operatorname{Re} x^*x_2 = 1$. It follows that $\{x : x \in X, \operatorname{Re} x^*x = 1\}$ is a support hyperplane for B_X at both x_1 and x_2. ■

Since a member x^* of the unit sphere of the dual space of a normed space X supports B_X at some point x of S_X if and only if $\operatorname{Re} x^*x = x^*x = 1$, the following result is essentially just a restatement of the preceding one.

5.1.16 Corollary. *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is rotund.*

(b) *No member x^* of S_{X^*} supports B_X at more than one point.*

(c) *For each x^* in S_{X^*}, there is no more than one x in S_X such that* $\operatorname{Re} x^*x = 1$.

(d) *For each x^* in S_{X^*}, there is no more than one x in S_X such that* $x^*x = 1$.

The next characterization of rotundity is important in approximation theory. It links rotundity to the possession by nonempty convex sets of a certain nearest point property.

5.1.17 Definition. A nonempty subset A of a metric space M is a *set of uniqueness* if, for every element x of M, there is no more than one element y of A such that $d(x, y) = d(x, A)$. The set A is a *set of existence* or *proximinal* if, for every element x of M, there is at least one element y of A such that $d(x, y) = d(x, A)$. The set A is a *Chebyshev set* if, for every element x of M, there is exactly one element y of A such that $d(x, y) = d(x, A)$; that is, if A is both a set of uniqueness and a set of existence.

5.1.18 Theorem. *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is rotund.*

(b) *Every nonempty convex subset of X is a set of uniqueness.*

(c) *Every nonempty closed convex subset of X is a set of uniqueness.*

PROOF. It will first be shown that (a) $\Rightarrow$ (b). Suppose that X is rotund, that C is a nonempty convex subset of X, and that $x_0 \in X$. The goal is to show that there are not two or more points of C closest to x_0. Since y is a point of C closest to x_0 if and only if $y - x_0$ is a point of $-x_0 + C$ closest to 0, it may be assumed that $x_0 = 0$. It may also be assumed that $d(0, C) > 0$ and then, after multiplying each point of C by the same positive constant, that $d(0, C) = 1$. Suppose that c_1 and c_2 are points of C closest to 0. Then $\|c_1\| = \|c_2\| = 1$, so

$$\{ tc_1 + (1 - t)c_2 : 0 \le t \le 1 \} \subseteq C \cap B_X \subseteq S_X.$$

Since X is rotund, it follows that $c_1 = c_2$, which proves that (a) $\Rightarrow$ (b).

It is clear that (b) $\Rightarrow$ (c), so all that remains to be shown is that (c) $\Rightarrow$ (a). Suppose that X is not rotund. Then there is a line segment in S_X of the form $\{ tx_1 + (1 - t)x_2 : 0 \le t \le 1 \}$, where $x_1 \neq x_2$. This line segment is a nonempty closed convex subset of X such that each of its infinitely many points is at the same distance from the origin. The failure of (a) therefore implies the failure of (c). ■

5.1.19 Corollary. (M. M. Day, 1941 [47]). *If a normed space is rotund and reflexive, then each of its nonempty closed convex subsets is a Chebyshev set.*

PROOF. Let C be a nonempty closed convex subset of a rotund reflexive normed space X and let x_0 be an element of X. Then there is a sequence (y_n) in C such that $\lim_n \|y_n - x_0\| = d(x_0, C)$. By the reflexivity of X, there is some subsequence (y_{n_j}) of the bounded sequence (y_n) that converges weakly to some y_0. Then $y_0 \in C$ since C is weakly closed, and y_0 is a point of C closest to x_0 since an application of Theorem 2.5.21 shows that

$$d(x_0, C) \leq \|y_0 - x_0\| \leq \liminf_j \|y_{n_j} - x_0\| = d(x_0, C).$$

Therefore there is at least one point of C closest to x_0. Notice that the rotundity of X has not yet been used.

Finally, the rotundity of X implies that C is a set of uniqueness, so there is no other point of C besides y_0 closest to x_0. Thus, the set C is Chebyshev. ■

It was shown in the proof of the preceding corollary that if a normed space X is reflexive, then each of its nonempty closed convex subsets is a set of existence. Conversely, it can be shown, as a straightforward corollary of James's theorem that appears in the optional Sections 1.13 and 2.9, that a *Banach* space is reflexive if every nonempty closed convex subset of the space is a set of existence, and a tiny bit of extra work based on a result of Jörg Blatter then shows that a *normed* space is reflexive if every nonempty closed convex subset of the space is a set of existence; see Exercise 5.11. Combining this with Theorem 5.1.18 and Corollary 5.1.19 shows that *a normed space is rotund and reflexive if and only if each of its nonempty closed convex subsets is a Chebyshev set.* This important result is sometimes called the Day-James theorem.

The last characterization of rotundity to be given here characterizes the property in terms of itself. One important ingredient in the characterization is the following obvious result.

5.1.20 Proposition. *If a normed space is rotund, then so is each of its subspaces.*

As a trivial consequence of this and the fact that every normed space is a subspace of itself, a normed space is rotund if and only if each of its subspaces is rotund. The following is a far better result along these lines.

5.1.21 Proposition. *A normed space is rotund if and only if each of its two-dimensional subspaces is rotund.*

PROOF. Suppose that X is a normed space that is not rotund. It is enough to find a two-dimensional subspace of X that is not rotund. Let x_1 and x_2 be distinct members of S_X such that $\frac{1}{2}(x_1 + x_2) \in S_X$. If there were a scalar α such that $x_1 = \alpha x_2$, then 1 and α would be different scalars of absolute value 1 such that $\frac{1}{2}(1+\alpha)$ also has absolute value 1, which would contradict the rotundity of $\mathbb{F}$. It follows that x_1 and x_2 are linearly independent. Therefore $\langle\{x_1, x_2\}\rangle$ is a two-dimensional subspace of X that is not rotund. ■

Proposition 5.1.20 leads naturally to the question of whether quotient spaces formed from rotund normed spaces must be rotund. Perhaps surprisingly, the answer is no, even for Banach spaces. The following example is from a 1959 paper by Victor Klee [137].

5.1.22 Example. This is a continuation of Example 5.1.8. The reader should review the following items from that example: property (3) of the functions f_m, the definition of the sets A_m and C, the fact that if $(\alpha_n) \in C$ then $|\alpha_n| \leq 1$ for each n, the fact that C is the closed unit ball of the rotund Banach space $\ell_{1,r}$, and the fact that $\|(\alpha_n)\|_r < 1$ whenever $(\alpha_n) \in \ell_{1,r}$ and $\sum_n f_{m(n)}\big(|\alpha_n|\big) < 1$.

Let Y be any separable nonrotund Banach space; for example, either c_0 or ℓ_1 will do. Let D be a countable dense subset of B_Y° and let g be a function from $\mathbb{N}$ onto D such that $g(A_m) \subseteq \frac{m}{m+1} B_Y$ for each m. Define $T\colon \ell_{1,r} \to Y$ by the formula $T\big((\alpha_n)\big) = \sum_n \alpha_n g(n)$; notice that the sum is absolutely convergent, so there is no problem with the definition. The function T is clearly linear.

Most of the remaining work in this example involves showing that T maps the open unit ball of $\ell_{1,r}$ onto the open unit ball of Y. To this end, it will first be shown that $T(C) \subseteq B_Y$. Suppose that $(\alpha_n) \in C$. Then

$$\begin{aligned}
\|T(\alpha_n)\| &= \Big\|\sum_n \alpha_n g(n)\Big\| \\
&\leq \sum_n |\alpha_n|\,\|g(n)\| \\
&\leq \sum_m \Big(\sum_{n\in A_m} |\alpha_n| \frac{m}{m+1}\Big) \\
&\leq \sum_m \Big(\sum_{n\in A_m} f_m(|\alpha_n|)\Big) \\
&= \sum_n f_{m(n)}\big(|\alpha_n|\big) \\
&\leq 1,
\end{aligned}$$

as claimed. Notice that this also shows that T is bounded, since $C = B_{\ell_{1,r}}$.

Let $U_{\ell_{1,r}}$ and U_Y be the open unit balls of $\ell_{1,r}$ and Y respectively. It will now be shown that $T(U_{\ell_{1,r}}) \supseteq U_Y$. Suppose that $y \in U_Y$. Since $\|2y\| < 2$, the open balls of radius 1 centered at $2y$ and the origin intersect (at y, for example), so there is a d_1 in D such that

$$\|2y - d_1\| < 1.$$

Since $\|4y - 2d_1\| < 2$, there is a d_2 in D such that

$$\|4y - 2d_1 - d_2\| < 1.$$

Since $\|8y - 4d_1 - 2d_2\| < 2$, there is a d_3 in D such that

$$\|8y - 4d_1 - 2d_2 - d_3\| < 1.$$

Continuing in the obvious fashion produces a sequence (d_n) in D such that

$$\left\| y - \sum_{n=1}^{k} 2^{-n} d_n \right\| < 2^{-k}$$

for each positive integer k, so $y = \sum_n 2^{-n} d_n$. Rearranging the terms of the sequence (2^{-n}) and interspersing zeros where necessary produces a member (γ_n) of $\ell_{1,r}$ such that $T(\gamma_n) = y$. It follows from property (3) of the functions f_m that $\sum_n f_{m(n)}(\gamma_n) < 1$, so $(\gamma_n) \in U_{\ell_{1,r}}$. This proves that $T(U_{\ell_{1,r}}) \supseteq U_Y$. An immediate consequence of this is that $T(\ell_{1,r}) = Y$, so T is an open mapping. Since $U_Y \subseteq T(U_{\ell_{1,r}}) \subseteq B_Y$ and $T(U_{\ell_{1,r}})$ is open, it must be that $T(U_{\ell_{1,r}}) = U_Y$.

By the first isomorphism theorem for Banach spaces, there is an isomorphism S from $\ell_{1,r}/\ker(T)$ onto Y. A glance at the proof of that theorem shows that S is the map from Theorem 1.7.13 such that $T = S\pi$, where π is the quotient map from $\ell_{1,r}$ onto $\ell_{1,r}/\ker(T)$. Let $U_{\ell_{1,r}/\ker(T)}$ be the open unit ball of $\ell_{1,r}/\ker(T)$. Then $\pi(U_{\ell_{1,r}}) = U_{\ell_{1,r}/\ker(T)}$ by Lemma 1.7.11. Since $T(U_{\ell_{1,r}}) = U_Y$, it must be that $S(U_{\ell_{1,r}/\ker(T)}) = U_Y$. It follows that S is an isometric isomorphism from $\ell_{1,r}/\ker(T)$ onto Y. Thus, the quotient space $\ell_{1,r}/\ker(T)$ of $\ell_{1,r}$ is not rotund even though $\ell_{1,r}$ is.

It does turn out that if X is a rotund normed space and M is a closed subspace of X that is a set of existence in X, then X/M is rotund. See Exercise 5.13.

Direct sums do a better job of preserving rotundity than do quotients.

5.1.23 Theorem. *Suppose that $X_1, \dots, X_n$ are normed spaces. Then $X_1 \oplus \cdots \oplus X_n$ is rotund if and only if each X_j is rotund.*

PROOF. This proof is based on the rotundity of real ℓ_2^2. It may be assumed that $n = 2$, for then an obvious induction argument based on the fact that $(X_1 \oplus \cdots \oplus X_{k-1}) \oplus X_k$ is isometrically isomorphic to $X_1 \oplus \cdots \oplus X_k$ when $2 \le k \le n$ yields the general case.

Suppose first that either X_1 or X_2 is not rotund. It follows from Proposition 1.8.10 (a) that $X_1 \oplus X_2$ has a subspace isometrically isomorphic to this nonrotund space, so $X_1 \oplus X_2$ is itself not rotund.

Now suppose conversely that both X_1 and X_2 are rotund. Let (x_1, x_2) and (y_1, y_2) be different members of $S_{X_1 \oplus X_2}$. The proof will be finished once it is shown that $\left\|\frac{1}{2}(x_1 + y_1, x_2 + y_2)\right\| < 1$. Notice that

$$\big(\|x_1\|, \|x_2\|\big), \big(\|y_1\|, \|y_2\|\big) \in S_{\ell_2^2}.$$

If either $\|x_1\| \neq \|y_1\|$ or $\|x_2\| \neq \|y_2\|$, then it follows from the rotundity of real ℓ_2^2 that

$$\begin{aligned}\left\|\tfrac{1}{2}(x_1 + y_1, x_2 + y_2)\right\| &= \left\|\tfrac{1}{2}\big(\|x_1 + y_1\|, \|x_2 + y_2\|\big)\right\|_2 \\ &\le \left\|\tfrac{1}{2}\big(\|x_1\| + \|y_1\|, \|x_2\| + \|y_2\|\big)\right\|_2 \\ &< 1,\end{aligned}$$

as required. Thus, it may be assumed that $\|x_1\| = \|y_1\|$ and $\|x_2\| = \|y_2\|$. It may also be assumed without loss of generality that $x_1 \neq y_1$. Then

$$\left\|\tfrac{1}{2}(x_1 + y_1)\right\| < \|x_1\| = \|y_1\| = \tfrac{1}{2}\big(\|x_1\| + \|y_1\|\big)$$

by the rotundity of X_1. Therefore

$$\begin{aligned}\left\|\tfrac{1}{2}(x_1 + y_1, x_2 + y_2)\right\| &= \left\|\tfrac{1}{2}\big(\|x_1 + y_1\|, \|x_2 + y_2\|\big)\right\|_2 \\ &< \left\|\tfrac{1}{2}\big(\|x_1\| + \|y_1\|, \|x_2\| + \|y_2\|\big)\right\|_2 \\ &= 1,\end{aligned}$$

which finishes the proof. ■

Much more could be said about basic rotundity, but for the purposes of this book it is time to move on to some of its special forms. The reader interested in learning more about rotundity in general should see Day's book [56, pp. 144–151] for an excellent brief discussion of it, and Istrăţescu's book [107] for a more thorough introduction to the subject.

Exercises

5.1 Suppose that (X_n) is a sequence of normed spaces and that $1 \le p < \infty$. Then the *ℓ_p sum* $\ell_p\big((X_n)\big)$ of the sequence (X_n) is the collection of all sequences (x_n) such that $x_n \in X_n$ for each n and $\sum_n \|x_n\|^p$ is finite, along with the obvious vector space operations and the norm defined by the formula

$$\|(x_n)\|_p = \Big(\sum_n \|x_n\|^p\Big)^{1/p}.$$

Prove that all of this really does define a normed space, and that $\ell_p\big((X_n)\big)$ is a Banach space if and only if each X_n is a Banach space.

5.2 With all notation as in the preceding exercise, suppose that $1 < p < \infty$. Prove that $\ell_p((X_n))$ is rotund if and only if each X_n is rotund. You may use the fact about the rotundity of certain Lebesgue spaces that is mentioned in Example 5.1.4.

5.3 Prove that if K is a compact Hausdorff space having more than one element, then $\text{rca}(K)$ is not rotund.

5.4 This exercise assumes some knowledge of inner product spaces. See, for example, [24] (in which such spaces are called pre-Hilbert spaces) or [202]. Prove that every inner product space is rotund.

5.5 Suppose that X is a normed space. Prove that the following are equivalent.

(a) The space X is rotund.

(b) Whenever (x_n) is a sequence in S_X and there is a member x^* of S_{X^*} such that $\lim_n x^*x_n = 1$, all weakly convergent subsequences of (x_n) have the same limit.

(c) Whenever (x_n) is a sequence in S_X and there is a member x^* of S_{X^*} such that $\lim_n x^*x_n = 1$, all convergent subsequences of (x_n) have the same limit.

5.6 (P. R. Beesack, E. Hughes, and M. Ortel, 1979 [22]). Suppose that X is a complex normed space. Prove that the following are equivalent.

(a) The space X is rotund.

(b) For each pair x_1 and x_2 of different members of S_X, there is a scalar α_{x_1,x_2} such that $\|\alpha_{x_1,x_2}x_1 + (1 - \alpha_{x_1,x_2})x_2\| < 1$.

5.7 Prove that a normed space is rotund if and only if no two closed balls in the space having disjoint interiors intersect at more than one point.

5.8 If C is a nonempty closed convex subset of a Hausdorff TVS X, then an element x_e of C is an *exposed point* of C if there is an x^* in X^* such that $\operatorname{Re} x^*$ is bounded from above on C and attains its supremum on C at x_e and only at x_e. (This definition was previously given in Exercise 2.112, in which it was shown that the property of being an exposed point is properly stronger than that of being an extreme point.) Prove that a normed space is rotund if and only if every point of its unit sphere is an exposed point of its closed unit ball.

5.9 Prove that if X is a Banach space that is not zero- or one-dimensional, then there is a nonrotund norm on X equivalent to the original norm.

5.10 Suppose that X is a normed space and that there is a one-to-one bounded linear operator T from X into a rotund normed space Y. Prove that the formula $\|x\|_r = \|x\| + \|Tx\|$ defines a rotund norm for X equivalent to its original norm $\|\cdot\|$.

5.11 This exercise requires James's theorem from either of the optional Sections 1.13 and 2.9.

(a) Prove that if every nonempty closed convex subset of a Banach space is a set of existence, then the space is reflexive.

(b) Prove that a Banach space is rotund and reflexive if and only if each of its nonempty closed convex subsets is a Chebyshev set.

(c) It is a 1976 result of Jörg Blatter [30] that a normed space is complete if every nonempty closed convex subset of the space has an element of minimum norm. Use this result to improve parts (a) and (b) of this exercise by replacing "Banach space" with "normed space."

5.12 (a) Show that every reflexive subspace of a normed space is a set of existence.

(b) Show that every nonempty weakly compact subset of a normed space is a set of existence.

(c) Show that every nonempty weakly* compact subset of the dual space of a normed space is a set of existence.

5.13 (a) Prove that if M is a subspace of a rotund normed space X and M is a set of existence in X, then X/M is rotund. (Notice that the fact that M is a set of existence implies that it is closed. Why?)

(b) From parts (a) of this exercise and Exercise 5.12, conclude that if M is a reflexive subspace of a rotund normed space X, then X/M is rotund. (This is a 1959 result of Victor Klee [137].)

5.2 Uniform Rotundity

By Proposition 5.1.2, a normed space is rotund if and only if every nontrivial straight line segment whose endpoints lie in the unit sphere of the space has its midpoint in the interior of the closed unit ball. This leads naturally to the question of how far into the interior of the closed unit ball the midpoint of such a segment must be if the segment has some minimum positive length, since this is a measure of the "amount of rotundity" the space has. As will be seen in Example 5.2.13, it is quite possible for a normed space X to be rotund and yet for there to be sequences (x_n) and (y_n) in S_X such that $\|x_n - y_n\|$ is bounded away from 0 even though $\sup_n \left\|\frac{1}{2}(x_n + y_n)\right\| = 1$. It seems reasonable to call a normed space X *uniformly rotund* if this does not happen; that is if, for every positive ϵ, there is a positive δ depending on ϵ such that $\left\|\frac{1}{2}(x+y)\right\| \le 1 - \delta$ whenever $x, y \in S_X$ and $\|x - y\| \ge \epsilon$.

5.2.1 Definition. (J. A. Clarkson, 1936 [41]). Let X be a normed space. Define a function $\delta_X\colon [0,2] \to [0,1]$ by the formula

$$\delta_X(\epsilon) = \inf\left\{\, 1 - \left\|\tfrac{1}{2}(x+y)\right\| : x, y \in S_X,\ \|x - y\| \ge \epsilon \,\right\}$$

if $X \ne \{0\}$, and by the formula

$$\delta_X(\epsilon) = \begin{cases} 0 & \text{if } \epsilon = 0; \\ 1 & \text{if } 0 < \epsilon \le 2 \end{cases}$$

if $X = \{0\}$. Then δ_X is the *modulus of rotundity* or *modulus of convexity* of X. The space X is *uniformly rotund* or *uniformly convex* if $\delta_X(\epsilon) > 0$ whenever $0 < \epsilon \le 2$.

It is easy to see that for each normed space X, the modulus of rotundity is a nondecreasing function of ϵ such that $\delta_X(0) = 0$. It is also clear that if M is a subspace of X, then $\delta_M(\epsilon) \ge \delta_X(\epsilon)$ when $0 \le \epsilon \le 2$. Furthermore, if X is a real one-dimensional normed space, then

$$\delta_X(\epsilon) = \begin{cases} 0 & \text{if } \epsilon = 0; \\ 1 & \text{if } 0 < \epsilon \le 2, \end{cases}$$

which means that the definition of $\delta_{\{0\}}$ given above is exactly what is needed to avoid having exceptions for the zero-dimensional case to the properties given at the beginning of this paragraph.

The first portion of this section is devoted to obtaining Theorem 5.2.5, which gives some alternative formulas for the modulus of rotundity that are often used as its definition. These formulas are not actually used in any crucial way later in this book. The main reason for developing them is to give some examples of a certain type of argument involving the geometry of two-dimensional real normed spaces that is often encountered in the theory of normed spaces but that has not yet been used in this book. A generous amount of detail will be provided to illustrate how much care must be taken to avoid making unwarranted assumptions about the relative positions of objects such as lines and points in the plane when carrying out such an argument, as well as other unjustified assumptions based on possibly unsupported intuition about the shape of closed unit balls.

5.2.2 Lemma. *Suppose that X is a two-dimensional real normed space. Then S_X is connected. Moreover, if $x^* \in X^*$, then $\{\, x : x \in S_X,\ x^*x \ge 0 \,\}$ is connected.*

PROOF. Only the second statement in the conclusion needs to be proved, since the first then follows by letting $x^* = 0$. By an easy argument involving the facts that X is isometrically isomorphic to a normed space with underlying vector space $\mathbb{R}^2$ and that continuous functions preserve connectedness, it may be assumed that the vector space underlying X is the same vector space $\mathbb{R}^2$ that underlies ℓ_2^2. Let $\|\cdot\|_X$ be the norm function of X and let x^* be a member of X^*. Define a map $f\colon \{\, x : x \in S_{\ell_2^2},\ x^*x \ge 0 \,\} \to \{\, x : x \in S_X, x^*x \ge 0 \,\}$ by the formula $f(x) = \|x\|_X^{-1}x$. It is easy to check that f is continuous and has range $\{\, x : x \in S_X, x^*x \ge 0 \,\}$. Since $\{\, x : x \in S_{\ell_2^2},\ x^*x \ge 0 \,\}$ is either a circle or a closed semicircle, it is connected, and therefore $\{\, x : x \in S_X,\ x^*x \ge 0 \,\}$ is also connected. ∎

5.2.3 Lemma. *Suppose that X is a normed space that is not zero-dimensional if $\mathbb{F} = \mathbb{C}$ and is neither zero- nor one-dimensional if $\mathbb{F} = \mathbb{R}$. If*

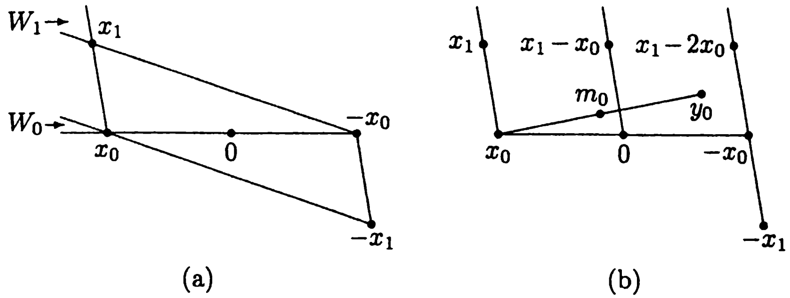


FIGURE 5.1. Two diagrams for the proof of Lemma 5.2.3.

$x_0 \in S_X$ and $y_0 \in B_X$, then there are members x_1 and y_1 of S_X such that $x_1 - y_1 = x_0 - y_0$ and $\left\|\frac{1}{2}(x_1 + y_1)\right\| \geq \left\|\frac{1}{2}(x_0 + y_0)\right\|$.

PROOF. Since X is a real normed space when multiplication of vectors by scalars is restricted to $\mathbb{R} \times X$, and since a one-dimensional complex normed space becomes a two-dimensional real normed space when treated this way, it may be assumed that $\mathbb{F} = \mathbb{R}$, and then that X is two-dimensional, and, for purposes of visualization, that the vector space underlying X is $\mathbb{R}^2$. It may also be assumed that $\|y_0\| < 1$.

Let x^* be a member of X^* whose kernel is the line through x_0 and $-x_0$ and such that $x^*y_0 \geq 0$. Then the set $\{ x : x \in S_X,\, x^*x \geq 0 \}$ is connected due to the preceding lemma. Since

$$\|x_0 + y_0 - x_0\| = \|y_0\| < 1$$

and

$$\|-x_0 + y_0 - x_0\| \geq 2\|x_0\| - \|y_0\| > 1,$$

the intermediate value theorem for connected sets assures that there is an x_1 in $\{ x : x \in S_X,\, x^*x \geq 0 \}$, necessarily different from both x_0 and $-x_0$, such that $\|x_1 + y_0 - x_0\| = 1$. Let $y_1 = x_1 + y_0 - x_0$. Then $x_1, y_1 \in S_X$ and $x_1 - y_1 = x_0 - y_0$, so all that remains to be proved is that $\left\|\frac{1}{2}(x_1 + y_1)\right\| \geq \left\|\frac{1}{2}(x_0 + y_0)\right\|$.

If $z_1 \in B_X$ and $z_2 \in B_X^\circ$, then $tz_1 + (1-t)z_2 \in B_X^\circ$ when $0 < t < 1$; this follows from the triangle inequality as well as from Lemma 2.2.18. This fact will be used repeatedly in the following arguments. Let $\overline{z_1, z_2}$ represent the straight line (not just a straight line segment) through the points z_1 and z_2 in X when $z_1 \neq z_2$. In Figure 5.1 (a), the set W_0 is the set of all points on or above $\overline{x_0, -x_0}$ and on or to the left of $\overline{x_0, -x_1}$, while W_1 is the set of all points strictly above $\overline{x_1, -x_0}$ and strictly to the left of $\overline{x_0, x_1}$. Notice that the parallelogram with vertices x_0, x_1, $-x_0$, and $-x_1$ lies in B_X by an easy convexity argument, and the interior of that parallelogram lies in B_X°. It has already been shown that, with the points x_0, $-x_0$, and x_1 arranged as they are in Figure 5.1 (a), the point y_0 lies on or above $\overline{x_0, -x_0}$. If y_0 were strictly to the left of $\overline{x_0, x_1}$, then it would follow that either $y_0 \in W_0$ or $y_1 \in W_1$. However, either case would place one of the two points x_0 and x_1 in the "interior" of a straight line segment with one endpoint in B_X and the other in B_X°, which would imply that either x_0 or x_1 is in B_X°. This contradiction shows that y_0 lies on or to the right of $\overline{x_0, x_1}$. If y_0 were to lie on this line, then an easy argument involving the location of y_1 would show that $\|x_1\| < 1$, another contradiction. Therefore y_0 lies strictly to the right of $\overline{x_0, x_1}$.

Since $-x_0$ and $-x_1$ are different elements of S_X, it follows that the points of $\overline{-x_0, -x_1}$ lying on or above $\overline{x_0, -x_0}$ all have norm at least 1, so no point on $\overline{x_0, y_0}$ strictly between x_0 and y_0 can intersect $\overline{-x_0, -x_1}$, nor can y_0 lie on that line; see Figure 5.1 (b). Thus, the point y_0 lies strictly to the left of $\overline{-x_0, -x_1}$. Since $\overline{x_0, x_1}$, $\overline{0, x_1 - x_0}$, and $\overline{-x_0, -x_1}$ are parallel, it also follows that $\frac{1}{2}(x_0+y_0)$, denoted by m_0 in Figure 5.1 (b), lies strictly between $\overline{x_0, x_1}$ and $\overline{0, x_1 - x_0}$. Therefore $\overline{0, \frac{1}{2}(x_0 + y_0)}$ intersects $\overline{x_0, x_1}$ somewhere on or above $\overline{x_0, -x_0}$ and at a point of $\overline{0, \frac{1}{2}(x_0 + y_0)}$ that is reached by traveling from 0 to $\frac{1}{2}(x_0 + y_0)$ and then beyond.

The point to the preceding argument is that there are real numbers s and t such that $s > 1$ and $t \geq 0$ that satisfy the equation

$$\frac{s}{2}(x_0 + y_0) = x_0 + t(x_1 - x_0).$$

It follows easily that

$$\frac{s}{2}(x_1 + y_1) = x_0 + (t + s)(x_1 - x_0).$$

If $t \leq 1$, then

$$\|x_0 + t(x_1 - x_0)\| \leq 1 \leq \|x_0 + (t + s)(x_1 - x_0)\|.$$

If $t > 1$, then the facts that $\|x_1\| = 1$, that $\|x_0 + (t + s)(x_1 - x_0)\| \geq 1$, and that $x_0 + t(x_1 - x_0)$ lies on the straight line segment connecting x_1 to $x_0 + (t + s)(x_1 - x_0)$ together assure that

$$\|x_0 + t(x_1 - x_0)\| \leq \|x_0 + (t + s)(x_1 - x_0)\|.$$

In either case, it follows that $\left\|\frac{1}{2}s(x_0 + y_0)\right\| \leq \left\|\frac{1}{2}s(x_1 + y_1)\right\|$, and therefore that $\left\|\frac{1}{2}(x_0 + y_0)\right\| \leq \left\|\frac{1}{2}(x_1 + y_1)\right\|$. ■

5.2.4 Lemma. *Suppose that X is a normed space, that $0 < \epsilon < 2$, and that x and y are members of S_X such that $\|x - y\| = \epsilon$. Then there are sequences (x_n) and (y_n) in S_X such that $\|x_n - y_n\| > \epsilon$ for each n and such that $\lim_n x_n = x$ and $\lim_n y_n = y$.*

PROOF. It may be assumed that X is a two-dimensional real normed space whose underlying vector space is $\mathbb{R}^2$. Let x^* be a nonzero member of X^*

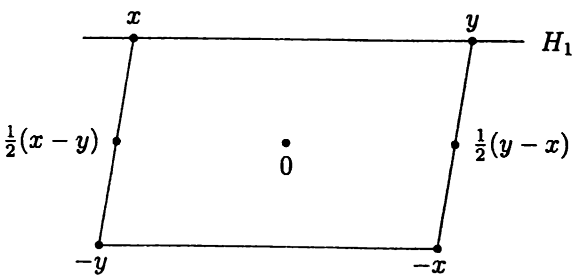


FIGURE 5.2. A diagram for the proof of Lemma 5.2.4.

such that $x^*(x-y)=0$. If $x^*x=0$, then it would follow that $x, y \in S_X \cap \ker x^*$ and therefore that $\|x-y\|$ is either 0 or 2, a contradiction. It may therefore be assumed that $x^*x = x^*y = 1$. For each real number δ, let $H_\delta = \{\, z : z \in X,\ x^*z = \delta \,\}$. The topmost horizontal line in Figure 5.2 represents H_1. The convexity of B_X assures that the parallelogram whose vertices are x, y, $-x$, and $-y$ lies in B_X. The conclusion of the lemma is immediate if the straight line segment $H_1 \cap B_X$ has length greater than ϵ, so it may be assumed that x and y are the endpoints of $H_1 \cap B_X$.

Suppose that $H_{\delta_0} \cap B_X$ were to have length ϵ for some δ_0 such that $0 < \delta_0 < 1$. Then some point of S_X would lie on the line through x and $-y$ somewhere strictly between those points, which would imply that $\|\frac{1}{2}(x-y)\| = 1$. Since $\|x-y\| = \epsilon < 2$, this is a contradiction. Therefore $H_\delta \cap B_X$ has length greater than ϵ when $0 < \delta < 1$. Let (δ_n) be a sequence of positive real numbers strictly increasing to 1. For each positive integer n, let x_n and y_n be the endpoints of $H_{\delta_n} \cap B_X$ such that x_n is on or to the left of the line through x and $-y$ in Figure 5.2 and y_n is on or to the right of the line through y and $-x$. Notice that $x_n, y_n \in S_X$ and $\|x_n - y_n\| > \epsilon$ for each n. Because of the compactness of S_X, it may be assumed that there are elements x_0 and y_0 of S_X such that $\lim_n x_n = x_0$ and $\lim_n y_n = y_0$. Since $x_0, y_0 \in H_1 \cap S_X$ and $\|x_0 - y_0\| \ge \epsilon$, it follows that $x = x_0$ and $y = y_0$. ■

5.2.5 Theorem. *Suppose that X is a normed space that is not zero-dimensional if $\mathbb{F} = \mathbb{C}$ and is neither zero- nor one-dimensional if $\mathbb{F} = \mathbb{R}$. Let δ_X be the modulus of rotundity of X. Then*

$$\begin{aligned}\delta_X(\epsilon) &= \inf\{\, 1 - \|\tfrac{1}{2}(x+y)\| : x, y \in B_X,\ \|x-y\| \ge \epsilon \,\}\\ &= \inf\{\, 1 - \|\tfrac{1}{2}(x+y)\| : x, y \in S_X,\ \|x-y\| = \epsilon \,\}\\ &= \inf\{\, 1 - \|\tfrac{1}{2}(x+y)\| : x, y \in B_X,\ \|x-y\| = \epsilon \,\}\end{aligned}$$

when $0 \le \epsilon \le 2$, and

$$\begin{aligned}\delta_X(\epsilon) &= \inf\{\, 1 - \|\tfrac{1}{2}(x+y)\| : x, y \in S_X,\ \|x-y\| > \epsilon \,\}\\ &= \inf\{\, 1 - \|\tfrac{1}{2}(x+y)\| : x, y \in B_X,\ \|x-y\| > \epsilon \,\}\end{aligned}$$

when $0 \le \epsilon < 2$.

PROOF. It may be assumed that $\mathbb{F} = \mathbb{R}$. Let δ_X be as in Definition 5.2.1 and let δ_1, δ_2, δ_3, δ_4, and δ_5 be the five infima in the statement of this theorem in the order in which they occur. The first claim is that $\delta_X(\epsilon) = \delta_1(\epsilon)$ when $0 \leq \epsilon \leq 2$. Suppose that $0 < \epsilon_0 \leq 2$, that $x_0, y_0 \in B_X$, and that $\|x_0 - y_0\| \geq \epsilon_0$. To prove the claim, it is enough to show that $\delta_X(\epsilon_0) \leq 1 - \left\|\frac{1}{2}(x_0 + y_0)\right\|$. It may be assumed that $x_0 \in S_X$. By Lemma 5.2.3, there are members x_1 and y_1 of S_X such that $\|x_1 - y_1\| = \|x_0 - y_0\| \geq \epsilon_0$ and $\left\|\frac{1}{2}(x_1 + y_1)\right\| \geq \left\|\frac{1}{2}(x_0 + y_0)\right\|$, so

$$\delta_X(\epsilon_0) \leq 1 - \left\|\tfrac{1}{2}(x_1 + y_1)\right\| \leq 1 - \left\|\tfrac{1}{2}(x_0 + y_0)\right\|.$$

Therefore $\delta_X(\epsilon) = \delta_1(\epsilon)$ when $0 \leq \epsilon \leq 2$, as claimed. A similar argument shows that $\delta_4(\epsilon) = \delta_5(\epsilon)$ when $0 \leq \epsilon < 2$.

The next claim is that $\delta_X(\epsilon) = \delta_4(\epsilon)$ when $0 \leq \epsilon < 2$. To show this, suppose that $0 < \epsilon_1 < 2$, that $x_2, y_2 \in S_X$, and that $\|x_2 - y_2\| = \epsilon_1$. It is enough to show that $\delta_4(\epsilon_1) \leq 1 - \left\|\frac{1}{2}(x_2 + y_2)\right\|$. By Lemma 5.2.4, there are sequences (v_n) and (w_n) in S_X such that $\|v_n - w_n\| > \epsilon_1$ for each n and such that $\lim_n v_n = x_2$ and $\lim_n w_n = y_2$. It follows that $\delta_4(\epsilon_1) \leq 1 - \left\|\frac{1}{2}(v_n + w_n)\right\|$ for each n, so passing to the limit yields the inequality needed to prove the claim.

It is next claimed that $\delta_2(\epsilon) = \delta_3(\epsilon)$ when $0 \leq \epsilon \leq 2$. Suppose that $0 < \epsilon_2 \leq 2$, that $x_3, y_3 \in B_X$, and that $\|x_3 - y_3\| = \epsilon_2$. To prove the claim, it suffices to show that $\delta_2(\epsilon_2) \leq 1 - \left\|\frac{1}{2}(x_3 + y_3)\right\|$. Suppose for the moment that x_3 and y_3 both have norm less than 1. Then there are two closed line segments of length ϵ_2 on the line through x_3 and y_3 such that each segment has one endpoint in S_X and the other in B_X°. Since $\frac{1}{2}(x_3 + y_3)$ is a convex combination of the midpoints of these two segments, the midpoint of at least one of the segments has norm at least $\left\|\frac{1}{2}(x_3 + y_3)\right\|$, so it may be assumed that $x_3 \in S_X$. Lemma 5.2.3 then produces elements x_4 and y_4 of S_X such that $\|x_4 - y_4\| = \|x_3 - y_3\| = \epsilon_2$ and $\left\|\frac{1}{2}(x_4 + y_4)\right\| \geq \left\|\frac{1}{2}(x_3 + y_3)\right\|$. Therefore

$$\delta_2(\epsilon_2) \leq 1 - \left\|\tfrac{1}{2}(x_4 + y_4)\right\| \leq 1 - \left\|\tfrac{1}{2}(x_3 + y_3)\right\|,$$

as required for the claim.

The final claim needed to prove the theorem is that $\delta_X(\epsilon) = \delta_2(\epsilon)$ when $0 \leq \epsilon \leq 2$. Suppose that $0 < \epsilon_3 < 2$ and that x_5 and y_5 are elements of S_X such that $\|x_5 - y_5\| \geq \epsilon_3$. It is enough to show that $\delta_2(\epsilon_3) \leq 1 - \left\|\frac{1}{2}(x_5 + y_5)\right\|$. The closed line segment with endpoints x_5 and y_5 includes a closed line segment of length exactly ϵ_3 centered at $\frac{1}{2}(x_5 + y_5)$, which by the argument of the preceding paragraph produces elements x_6 and y_6 of S_X such that $\|x_6 - y_6\| = \epsilon_3$ and $\left\|\frac{1}{2}(x_6 + y_6)\right\| \geq \left\|\frac{1}{2}(x_5 + y_5)\right\|$. Therefore

$$\delta_2(\epsilon_3) \leq 1 - \left\|\tfrac{1}{2}(x_6 + y_6)\right\| \leq 1 - \left\|\tfrac{1}{2}(x_5 + y_5)\right\|,$$

which finishes the proof of the claim and of the theorem. ■

It is time to leave the general study of the modulus of rotundity and to focus on uniform rotundity. Notice that if X is a uniformly rotund normed space with modulus of rotundity δ_X and x and y are different elements of S_X, then $\left\|\frac{1}{2}(x+y)\right\| \leq 1 - \delta_X\big(\|x-y\|\big) < 1$, which proves the most basic fact about uniformly rotund normed spaces.

5.2.6 Proposition. *Every uniformly rotund normed space is rotund.*

As is true for rotundity, uniform rotundity is not always preserved by isomorphisms. For example, real Euclidean 2-space is clearly uniformly rotund even though it is isomorphic to real ℓ_1^2, a nonrotund space. The following statement, whose proof is obvious, is about all that can be said in this direction.

5.2.7 Proposition. *Every normed space that is isometrically isomorphic to a uniformly rotund normed space is itself uniformly rotund.*

Though the definition of uniform rotundity allows its nature to be visualized easily, the following sequential characterizations of the property are sometimes easier to use in applications.

5.2.8 Proposition. *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is uniformly rotund.*
(b) *Whenever (x_n) and (y_n) are sequences in S_X and $\left\|\frac{1}{2}(x_n+y_n)\right\| \to 1$, it follows that $\|x_n - y_n\| \to 0$.*
(c) *Whenever (x_n) and (y_n) are sequences in B_X and $\left\|\frac{1}{2}(x_n+y_n)\right\| \to 1$, it follows that $\|x_n - y_n\| \to 0$.*
(d) *Whenever (x_n) and (y_n) are sequences in X and $\|x_n\|$, $\|y_n\|$, and $\left\|\frac{1}{2}(x_n+y_n)\right\|$ all tend to 1, it follows that $\|x_n - y_n\| \to 0$.*

PROOF. Suppose that (b) holds and that (x_n) and (y_n) are sequences in X such that $\|x_n\| \to 1$, $\|y_n\| \to 1$, and $\left\|\frac{1}{2}(x_n+y_n)\right\| \to 1$. It will be shown that $\|x_n - y_n\| \to 0$. By discarding terms from the beginning of the sequences if necessary, it may be assumed that no x_n or y_n is 0. Then

$$\begin{aligned} 1 &\geq \left\|\tfrac{1}{2}\big(\|x_n\|^{-1}x_n + \|y_n\|^{-1}y_n\big)\right\| \\ &\geq \left\|\tfrac{1}{2}(x_n+y_n)\right\| - \left\|\tfrac{1}{2}(1-\|x_n\|^{-1})x_n\right\| - \left\|\tfrac{1}{2}(1-\|y_n\|^{-1})y_n\right\| \\ &\to 1, \end{aligned}$$

so $\left\|\frac{1}{2}\big(\|x_n\|^{-1}x_n + \|y_n\|^{-1}y_n\big)\right\| \to 1$. Since $\left\|\,\|x_n\|^{-1}x_n - \|y_n\|^{-1}y_n\right\| \to 0$ by (b), it follows that

$$\begin{aligned} 0 &\leq \|x_n - y_n\| \\ &\leq \left\|\,\|x_n\|^{-1}x_n - \|y_n\|^{-1}y_n\right\| + \left\|(1-\|x_n\|^{-1})x_n\right\| + \left\|(1-\|y_n\|^{-1})y_n\right\| \\ &\to 0. \end{aligned}$$

Therefore $\|x_n - y_n\| \to 0$, which establishes that (b) $\Rightarrow$ (d).

Now suppose that (d) holds and that (x_n) and (y_n) are sequences in B_X such that $\left\|\frac{1}{2}(x_n + y_n)\right\| \to 1$. Since $\left\|\frac{1}{2}(x_n + y_n)\right\| \le \frac{1}{2}\big(\|x_n\| + \|y_n\|\big)$ for each n, it follows that $\|x_n\| \to 1$ and $\|y_n\| \to 1$, so $\|x_n - y_n\| \to 0$ by (d). Therefore (d) $\Rightarrow$ (c), from which it follows that (b) $\Leftrightarrow$ (c) $\Leftrightarrow$ (d).

Suppose that X is uniformly rotund and that (x_n) and (y_n) are sequences in S_X such that $\left\|\frac{1}{2}(x_n + y_n)\right\| \to 1$ but $\|x_n - y_n\|$ does not tend to 0. Let δ_X be the modulus of rotundity of X. It follows that there is a subsequence (x_{n_j}) of (x_n) such that $\|x_{n_j} - y_{n_j}\| \ge \epsilon$ for some positive ϵ and each j, which implies that $\left\|\frac{1}{2}(x_{n_j} + y_{n_j})\right\| \le 1 - \delta_X(\epsilon)$ for each j, a contradiction. Therefore (a) $\Rightarrow$ (b).

Finally, suppose that X is not uniformly rotund. Then there is an ϵ such that $0 < \epsilon \le 2$ but $\delta_X(\epsilon) = 0$. Therefore there are sequences (x_n) and (y_n) in S_X such that $\|x_n - y_n\| \ge \epsilon$ for each n but $\left\|\frac{1}{2}(x_n + y_n)\right\| \to 1$, so (b) does not hold. This shows that (b) $\Rightarrow$ (a). ■

One very important class of uniformly rotund Banach spaces is that of the spaces $L_p(\Omega, \Sigma, \mu)$, where μ is a positive measure on a σ-algebra Σ of subsets of a set Ω and $1 < p < \infty$; notice that this includes the spaces ℓ_p and ℓ_p^n such that $1 < p < \infty$ and n is a nonnegative integer. The uniform rotundity of these spaces was first established in 1936 by James Clarkson [41] in the same paper in which he introduced the notions of rotundity and uniform rotundity. Proofs of this usually begin by establishing some inequality or inequalities in the scalar field or for the norms of elements of $L_p(\Omega, \Sigma, \mu)$, and often require the separate consideration of the cases in which $1 < p \le 2$ and $2 \le p < \infty$; see, for example, Clarkson's paper or [21]. The proof given here is from [141] and is based on a method from a 1950 paper by E. J. McShane [166] that does not require the consideration of the two different cases.

The following lemma is proved in quite a bit more generality than is needed for the proof of Clarkson's result, but the extra generality will be required later for the proof of Theorem 5.2.25.

5.2.9 Lemma. *For each p such that $1 < p < \infty$ and each function $\lambda\colon (0,2] \to (0,1]$, there is a function $\gamma_{p,\lambda}\colon (0,2] \to (0,1]$ such that if X is a uniformly rotund normed space whose modulus of rotundity δ_X has the property that $\lambda(\epsilon) \le \delta_X(\epsilon)$ when $0 < \epsilon \le 2$, then*

$$\left\|\tfrac{1}{2}(x+y)\right\|^p \le \big(1 - \gamma_{p,\lambda}(t)\big)\left(\frac{\|x\|^p + \|y\|^p}{2}\right)$$

whenever $0 < t \le 2$ and x and y are members of X such that $\|x - y\| \ge t\max\{\|x\|, \|y\|\}$.

Proof. Suppose to the contrary that there were a p such that $1 < p < \infty$ and a function $\lambda\colon (0,2] \to (0,1]$ for which no such function $\gamma_{p,\lambda}$ exists.

Let $f(t) = \left(\frac{1}{2}(1+t)\right)^p / \left(\frac{1}{2}(1+t^p)\right)$ when $0 \le t \le 1$. It is an easy calculus exercise to show that f strictly increases on $[0,1]$ to its maximum value of 1, from which it follows that if X is a normed space and x and y are members of X such that $\|x\| = 1$ and $\|y\| \le 1$, then

$$\frac{\left\|\frac{1}{2}(x+y)\right\|^p}{\frac{1}{2}\left(\|x\|^p + \|y\|^p\right)} \le \frac{\left(\frac{1}{2}(1+\|y\|)\right)^p}{\frac{1}{2}\left(1+\|y\|^p\right)} \le 1.$$

By the supposition of the nonexistence of a function $\gamma_{p,\lambda}$ with the required properties, there must be a t such that $0 < t \le 2$, a sequence of uniformly rotund normed spaces (X_n) such that the modulus of rotundity δ_{X_n} of each X_n has the property that $\lambda(\epsilon) \le \delta_{X_n}(\epsilon)$ when $0 < \epsilon \le 2$, and sequences (x_n) and (y_n) such that for each positive integer n,

(1) $x_n, y_n \in X_n$;

(2) $\|x_n\| = 1$ and $\|y_n\| \le 1$;

(3) $\|x_n - y_n\| \ge t\max\{\|x_n\|, \|y_n\|\} = t$; and

(4) $\left\|\frac{1}{2}(x_n + y_n)\right\|^p > \left(1 - \dfrac{1}{n}\right)\left(\dfrac{\|x_n\|^p + \|y_n\|^p}{2}\right)$.

Notice that

(5) $\displaystyle\lim_n \frac{\left\|\frac{1}{2}(x_n + y_n)\right\|^p}{\frac{1}{2}\left(\|x_n\|^p + \|y_n\|^p\right)} = 1.$

It is an easy consequence of what has been proved about f that $\|y_n\| \to 1$, so in particular it may be assumed that no y_n is zero. Let $z_n = \|y_n\|^{-1} y_n$ for each n. Then $\|z_n - y_n\| \to 0$, so it may be assumed that $\|x_n - z_n\| \ge \frac{1}{2}t$ for each n. It follows that

$$\left\|\tfrac{1}{2}(x_n + z_n)\right\| \le 1 - \delta_{X_n}\left(\tfrac{1}{2}t\right) \le 1 - \lambda\left(\tfrac{1}{2}t\right) < 1$$

for each n. However, it also follows from (5) that

$$\lim_n \left\|\tfrac{1}{2}(x_n + z_n)\right\| = \lim_n \left\|\tfrac{1}{2}(x_n + y_n)\right\| = 1,$$

a contradiction. ■

With all notation as in the statement of the preceding lemma, letting λ be the restriction of the modulus of rotundity of the scalar field to $(0,2]$ produces the following special case of the lemma, which is what is actually needed for the proof of Theorem 5.2.11.

5.2.10 Lemma. *Suppose that $1 < p < \infty$. Then there is a function $\gamma_p\colon (0,2] \to (0,1]$ such that*

$$\left|\frac{\alpha+\beta}{2}\right|^p \le \left(1 - \gamma_p(t)\right)\left(\frac{|\alpha|^p + |\beta|^p}{2}\right)$$

when $0 < t \le 2$ and α and β are scalars such that $|\alpha - \beta| \ge t\max\{|\alpha|, |\beta|\}$.

5.2.11 Theorem. (J. A. Clarkson, 1936 [41]). *Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω and that $1 < p < \infty$. Then $L_p(\Omega, \Sigma, \mu)$ is uniformly rotund.*

PROOF. Suppose that $f_1, f_2 \in S_{L_p(\Omega,\Sigma,\mu)}$ and that $\|f_1 - f_2\|_p \geq \epsilon > 0$. Let

$$A = \left\{ \omega : \omega \in \Omega,\ |f_1(\omega) - f_2(\omega)|^p \geq \frac{\epsilon^p}{4}(|f_1(\omega)|^p + |f_2(\omega)|^p) \right\},$$

and observe that $|f_1(\omega) - f_2(\omega)| \geq (\epsilon/4^{1/p}) \max\{|f_1(\omega)|, |f_2(\omega)|\}$ when $\omega \in A$. With γ_p as in Lemma 5.2.10, it follows from that lemma that

$$\left|\frac{f_1(\omega) + f_2(\omega)}{2}\right|^p \leq \left(1 - \gamma_p\left(\frac{\epsilon}{4^{1/p}}\right)\right)\left(\frac{|f_1(\omega)|^p + |f_2(\omega)|^p}{2}\right)$$

whenever $\omega \in A$ and that

$$\left|\frac{f_1(\omega) + f_2(\omega)}{2}\right|^p \leq \frac{|f_1(\omega)|^p + |f_2(\omega)|^p}{2}$$

whenever $\omega \in \Omega$, so

$$\begin{aligned} 1 - \left\|\tfrac{1}{2}(f_1 + f_2)\right\|_p^p &= \int_\Omega \left(\frac{|f_1|^p + |f_2|^p}{2} - \left|\frac{f_1 + f_2}{2}\right|^p\right) d\mu \\ &\geq \int_A \left(\frac{|f_1|^p + |f_2|^p}{2} - \left|\frac{f_1 + f_2}{2}\right|^p\right) d\mu \\ &\geq \gamma_p\left(\frac{\epsilon}{4^{1/p}}\right) \int_A \frac{|f_1|^p + |f_2|^p}{2}\, d\mu. \end{aligned}$$

Let $\mathbf{I}_A$ be the indicator function of A. Then

$$\begin{aligned} \|f_1\mathbf{I}_A - f_2\mathbf{I}_A\|_p^p &= \|f_1 - f_2\|_p^p - \int_{\Omega\setminus A} |f_1 - f_2|^p\, d\mu \\ &\geq \epsilon^p - \frac{\epsilon^p}{4}\int_{\Omega\setminus A} (|f_1|^p + |f_2|^p)\, d\mu \\ &\geq \epsilon^p - \frac{\epsilon^p}{4}(\|f_1\|_p^p + \|f_2\|_p^p) \\ &= \frac{\epsilon^p}{2}, \end{aligned}$$

from which it follows that $\max\{\|f_1\mathbf{I}_A\|_p, \|f_2\mathbf{I}_A\|_p\} \geq \epsilon/(2 \cdot 2^{1/p})$. Therefore

$$1 - \left\|\tfrac{1}{2}(f_1 + f_2)\right\|_p^p \geq \gamma_p\left(\frac{\epsilon}{4^{1/p}}\right) \frac{\|f_1\mathbf{I}_A\|_p^p + \|f_2\mathbf{I}_A\|_p^p}{2} \geq \gamma_p\left(\frac{\epsilon}{4^{1/p}}\right)\frac{\epsilon^p}{2^{p+2}},$$

and so

$$\left\|\tfrac{1}{2}(f_1 + f_2)\right\|_p \leq \left(1 - \gamma_p\left(\frac{\epsilon}{4^{1/p}}\right)\frac{\epsilon^p}{2^{p+2}}\right)^{1/p} < 1.$$

The uniform rotundity of $L_p(\Omega, \Sigma, \mu)$ follows immediately. ■

5.2.12 Corollary. *Suppose that $1 < p < \infty$. Then ℓ_p is uniformly rotund, as is ℓ_p^n whenever n is a nonnegative integer.*

5.2.13 Example. This is an example of a rotund Banach space that is not uniformly rotund. For each positive integer n, let $p_n = 1 + \frac{1}{n}$. The example is based on an observation about the spaces $\ell_{p_n}^2$, each of which is itself uniformly rotund by Corollary 5.2.12. It is easy to check that $(1,0),(0,1) \in S_{\ell_{p_n}^2}$, that $\|(1,0)-(0,1)\|_{p_n} = 2^{n/(n+1)} \geq \sqrt{2}$, and that $\left\|\frac{1}{2}((1,0)+(0,1))\right\|_{p_n} = 2^{-1/(n+1)}$ for each n. Since $\lim_n 2^{-1/(n+1)} = 1$, the spaces $\ell_{p_n}^2$ such that $n \in \mathbb{N}$ are not "uniformly" uniformly rotund. In particular, *if a normed space X has each space $\ell_{p_n}^2$ such that $n \in \mathbb{N}$ isometrically embedded in it, then X cannot be uniformly rotund.*

Let X be the collection of all sequences (x_n) such that $x_n \in \ell_{p_n}^2$ for each n and $\sum_n \|x_n\|_{p_n}^2$ is finite, along with the obvious vector space operations and the norm given by the formula

$$\|(x_n)\|_2 = \Big(\sum_n \|x_n\|_{p_n}^2\Big)^{1/2}.$$

It is not difficult to check that this defines a Banach space; see Exercise 5.1. Suppose that $(x_{n,1}), (x_{n,2}), \frac{1}{2}((x_{n,1})+(x_{n,2})) \in S_X$. Then the triangle inequalities for ℓ_2 and each $\ell_{p_n}^2$ imply that

$$\begin{aligned}
1 &= \frac{1}{2}\Big(\sum_n \|x_{n,1}+x_{n,2}\|_{p_n}^2\Big)^{1/2} \\
&\leq \frac{1}{2}\Big(\sum_n \big(\|x_{n,1}\|_{p_n} + \|x_{n,2}\|_{p_n}\big)^2\Big)^{1/2} \\
&\leq \frac{1}{2}\Big(\sum_n \|x_{n,1}\|_{p_n}^2\Big)^{1/2} + \frac{1}{2}\Big(\sum_n \|x_{n,2}\|_{p_n}^2\Big)^{1/2} \\
&= 1,
\end{aligned}$$

so each of the two inequality symbols actually represents an equality. It follows that $(\|x_{n,1}\|_{p_n}), (\|x_{n,2}\|_{p_n}), \frac{1}{2}\Big((\|x_{n,1}\|_{p_n}) + (\|x_{n,2}\|_{p_n})\Big) \in S_{\ell_2}$, so the rotundity of ℓ_2 implies that $\|x_{n,1}\|_{p_n} = \|x_{n,2}\|_{p_n}$ for each n. It also follows that $\|x_{n,1}+x_{n,2}\|_{p_n} = \|x_{n,1}\|_{p_n} + \|x_{n,2}\|_{p_n}$ for each n. Therefore $x_{n,1} = x_{n,2}$ for every n by the rotundity of each $\ell_{p_n}^2$, which shows that $(x_{n,1}) = (x_{n,2})$ and thus that X is rotund.

Fix a positive integer m. To show that X is not uniformly rotund, it is enough to produce a subspace of X isometrically isomorphic to $\ell_{p_m}^2$. Letting

$$X_m = \{\, (x_n) : (x_n) \in X,\ x_n = 0 \text{ when } n \neq m \,\}$$

does precisely this.

It is clear that the Banach space of the preceding example is infinite-dimensional. As the next result shows, this is an essential property of any rotund normed space that is not uniformly rotund.

5.2.14 Proposition. *A finite-dimensional normed space is uniformly rotund if and only if it is rotund.*

PROOF. Suppose that X is a finite-dimensional rotund normed space. It is enough to show that X is uniformly rotund. Let (x_n) and (y_n) be sequences in S_X such that $\|x_n - y_n\|$ does not tend to 0. It is enough to show that $\|\frac{1}{2}(x_n + y_n)\|$ does not tend to 1. After thinning the sequences, it may be assumed that there is a positive ϵ such that $\|x_n - y_n\| \geq \epsilon$ for each n and that there are elements x and y of S_X such that $x_n \to x$ and $y_n \to y$. Then $\|x - y\| \geq \epsilon$, and $\|\frac{1}{2}(x_n + y_n)\| \to \|\frac{1}{2}(x + y)\| < 1$. ■

As was seen in Example 5.1.8, rotund Banach spaces do not have to be reflexive. It turns out that uniformly rotund Banach spaces do.

5.2.15 The Milman-Pettis Theorem. (D. P. Milman, 1938 [169]; B. J. Pettis, 1939 [183]). *Every uniformly rotund Banach space is reflexive.*

PROOF. This proof is due to Lindenstrauss and Tzafriri [157, p. 61]. Suppose that X is a uniformly rotund Banach space, which may be assumed to be infinite-dimensional, and that $x^{**} \in S_{X^{**}}$. Let Q be the natural map from X into X^{**}. By Goldstine's theorem, there is a net $(x_\alpha)_{\alpha \in I}$ in B_X such that $Qx_\alpha \xrightarrow{w^*} x^{**}$. Declaring that $(\alpha_1, \beta_1) \preceq (\alpha_2, \beta_2)$ when $\alpha_1 \preceq \alpha_2$ and $\beta_1 \preceq \beta_2$ then makes $\big(Q(\frac{1}{2}(x_\alpha + x_\beta))\big)_{(\alpha,\beta) \in I \times I}$ into a net. An easy argument based on Theorem 2.2.9 (g) shows that $Q(\frac{1}{2}(x_\alpha + x_\beta)) \xrightarrow{w^*} x^{**}$, so $\|\frac{1}{2}(x_\alpha + x_\beta)\| \to 1$ by Theorem 2.6.14. It then follows from the uniform rotundity of X that $\|x_\alpha - x_\beta\| \to 0$, which in turn implies that (x_α) is a Cauchy net in X. By Proposition 2.1.49, the completeness of X implies that (x_α) has some limit x_0, so $Qx_\alpha \to Qx_0$. Therefore $x^{**} = Qx_0$, so X is reflexive. ■

5.2.16 Corollary. *If a normed space X is isomorphic to a uniformly rotund Banach space, then X is reflexive.*

Notice that if a property P defined for normed spaces is such that whenever a normed space X has it then so do all normed spaces isometrically isomorphic to X, then a normed space $(X, \|\cdot\|_X)$ is isomorphic to some normed space $(Y, \|\cdot\|_Y)$ with property P if and only if there is a norm $\|\cdot\|_P$ for X equivalent to $\|\cdot\|_X$ such that $(X, \|\cdot\|_P)$ has property P. The forward implication comes about by letting $\|x\|_P = \|Tx\|_Y$ for each x in X, where T is the isomorphism, while the converse is obtained by noting that the identity map on X is an isomorphism from $(X, \|\cdot\|_X)$ onto $(X, \|\cdot\|_P)$. Thus, the preceding corollary can be viewed as the statement that every normed space with an equivalent uniformly rotund Banach norm is reflexive.

A normed space is called *superreflexive* if it can be given a uniformly rotund norm equivalent to its original norm (though that is not the original definition of superreflexivity; see [56, pp. 168–173] for the original definition and a proof that it is equivalent to the one given here). It follows from Corollary 5.2.16 that every superreflexive Banach space is reflexive. Mahlon Day showed in a 1941 paper [47] that there are reflexive spaces that are not superreflexive.

In a way, the Milman-Pettis theorem explains why it should not be surprising that the classical nonreflexive Banach spaces examined in the introductory examples of Section 5.1 are all nonrotund, and why it took so much work to construct a nonreflexive rotund Banach space in Example 5.1.8. A nonreflexive rotund Banach space X cannot be uniformly rotund, which means that while the midpoint of every straight line segment connecting distinct points on the unit sphere of X must dip into the interior of the closed unit ball, there is some positive ϵ such that one can always find two points on the unit sphere at least distance ϵ apart with the midpoint of the line segment connecting the points taking as shallow a dip into B_X° as one might wish. Such a unit sphere is somewhat oddly shaped.

Incidentally, it follows from the Milman-Pettis theorem that the rotund nonreflexive Banach space $\ell_{1,r}$ of Example 5.1.8 is another example of a rotund Banach space that is not uniformly rotund.

By Corollary 5.1.19, every nonempty closed convex subset of a rotund reflexive normed space is a Chebyshev set, so the following result is an immediate consequence of the Milman-Pettis theorem.

5.2.17 Corollary. *Every nonempty closed convex subset of a uniformly rotund Banach space is a Chebyshev set.*

Recall that a normed space has the *Radon-Riesz property* if, whenever (x_n) is a sequence in the space and x an element of the space such that $x_n \xrightarrow{w} x$ and $\|x_n\| \to \|x\|$, it follows that $x_n \to x$. As was discussed in Section 2.5, J. Radon and F. Riesz proved that the Lebesgue spaces $L_p(\Omega, \Sigma, \mu)$ such that $1 < p < \infty$ have this property. It turns out that this is a direct consequence of the uniform rotundity of those spaces.

5.2.18 Theorem. *Every uniformly rotund normed space has the Radon-Riesz property.*

PROOF. Suppose that X is a uniformly rotund normed space and that (x_n) is a sequence in X and x an element of X such that $x_n \xrightarrow{w} x$ and $\|x_n\| \to \|x\|$. It is to be shown that $x_n \to x$, so it may be assumed that $x \neq 0$ and therefore, after discarding some initial terms of (x_n) if necessary, that each x_n is nonzero. Now $\|x_n\|^{-1}x_n \xrightarrow{w} \|x\|^{-1}x$, and it is enough to show that $\|x_n\|^{-1}x_n \to \|x\|^{-1}x$, so it may even be assumed that x and each x_n are in S_X. Since $\frac{1}{2}(x_n+x) \xrightarrow{w} x$, the weak lower semicontinuity of the norm function implies that $\left\|\frac{1}{2}(x_n+x)\right\| \to 1$, so $\|x_n - x\| \to 0$, as required. ■

5.2.19 Corollary. (J. Radon, 1913 [192]; F. Riesz, 1928–1929 [196, 197]). *Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω and that $1 < p < \infty$. Then $L_p(\Omega, \Sigma, \mu)$ has the Radon-Riesz property.*

Like rotundity, uniform rotundity is inherited by subspaces, as can be verified by a glance at Proposition 5.2.8. A slightly better proof of this fact uses an important, though obvious, relationship between the modulus of rotundity of a normed space and those of its subspaces. That relationship is given in the next lemma, which was previously stated as an observation after Definition 5.2.1. The proposition following it is then an immediate consequence of the definition of uniform rotundity.

5.2.20 Lemma. *Suppose that M is a subspace of a normed space X and that δ_M and δ_X are the respective moduli of rotundity of the two spaces. Then $\delta_M(\epsilon) \geq \delta_X(\epsilon)$ when $0 \leq \epsilon \leq 2$.*

5.2.21 Proposition. *Every subspace of a uniformly rotund normed space is uniformly rotund.*

As was shown in Example 5.1.22, it is possible for a quotient space of a rotund normed space not to be rotund. In contrast, uniform rotundity is always inherited by quotient spaces. Two proofs of this will be given, one that is based on the following lemma and one that is not. Notice that, in the notation of this lemma, it is *not* claimed that $\delta_{X/M}(2) \geq \delta_X(2)$. The reason for this will be made clear in the example that follows the lemma.

5.2.22 Lemma. *Suppose that M is a closed subspace of a normed space X and that $\delta_{X/M}$ and δ_X are the moduli of rotundity of X/M and X respectively. Then $\delta_{X/M}(\epsilon) \geq \delta_X(\epsilon)$ when $0 \leq \epsilon < 2$.*

PROOF. It may be assumed that X/M is not zero-dimensional if $\mathbb{F} = \mathbb{C}$ and is neither zero- nor one-dimensional if $\mathbb{F} = \mathbb{R}$. Suppose that $x+M, y+M \in S_{X/M}$ and that $\|(x-y)+M\| > \epsilon$. By Theorem 5.2.5, it is enough to show that $1 - \left\|\frac{1}{2}(x+y)+M\right\| \geq \delta_X(\epsilon)$. It follows from Proposition 1.7.6 that there are sequences (x_n) and (y_n) in X such that $\|x_n\| \geq 1$, $\|y_n\| \geq 1$, $x_n + M = x + M$, and $y_n + M = y + M$ for each n, and $\|x_n\| \to 1$ and $\|y_n\| \to 1$. Then

$$\begin{aligned}\big\| \|x_n\|^{-1}x_n - \|y_n\|^{-1}y_n \big\| &\geq \|x_n - y_n\| - (1 - \|x_n\|^{-1})\|x_n\| - (1 - \|y_n\|^{-1})\|y_n\| \\ &\geq \|(x-y)+M\| - (\|x_n\| - 1) - (\|y_n\| - 1)\end{aligned}$$

for each n, so it may be assumed that $\big\| \|x_n\|^{-1}x_n - \|y_n\|^{-1}y_n \big\| > \epsilon$ for each n. Therefore

$$\begin{aligned}
\delta_X(\epsilon) &\le 1 - \big\| \tfrac{1}{2}\big(\|x_n\|^{-1}x_n + \|y_n\|^{-1}y_n\big) \big\| \\
&\le 1 - \big\| \tfrac{1}{2}\big(\|x_n\|^{-1}x_n + \|y_n\|^{-1}y_n\big) + M \big\| \\
&\le 1 - \big\| \tfrac{1}{2}(x_n + y_n) + M \big\| \\
&\qquad + \tfrac{1}{2}(1 - \|x_n\|^{-1})\|x_n + M\| + \tfrac{1}{2}(1 - \|y_n\|^{-1})\|y_n + M\| \\
&= 1 - \big\| \tfrac{1}{2}(x + y) + M \big\| + \tfrac{1}{2}(1 - \|x_n\|^{-1}) + \tfrac{1}{2}(1 - \|y_n\|^{-1})
\end{aligned}$$

for each n, so letting n tend to infinity produces the inequality needed to prove the lemma. ■

5.2.23 Example. Let $\ell_{1,r}$ be the rotund nonreflexive Banach space of Examples 5.1.8 and 5.1.22 and let $\delta_{\ell_{1,r}}$ be its modulus of rotundity. It follows readily from the rotundity of $\ell_{1,r}$ that $\delta_{\ell_{1,r}}(2) = 1$; see Exercise 5.14. It was shown in Example 5.1.22 that if Y is a separable nonrotund Banach space, then there is a closed subspace M of $\ell_{1,r}$ such that $\ell_{1,r}/M$ is isometrically isomorphic to Y, so in particular there is a closed subspace M_0 of $\ell_{1,r}$ such that $\ell_{1,r}/M_0$ is isometrically isomorphic to ℓ_∞^2. Then the modulus of rotundity $\delta_{\ell_{1,r}/M_0}$ for this quotient space is the same as that for ℓ_∞^2, so it follows easily that $\delta_{\ell_{1,r}/M_0}(2) = 0$. Therefore $\delta_{\ell_{1,r}}(2) > \delta_{\ell_{1,r}/M_0}(2)$, which shows why the preceding lemma was stated and proved only proved for values of ϵ strictly less than 2.

Incidentally, this example relies on the fact that $\ell_{1,r}$ is not reflexive. See Exercise 5.19 for the reason.

5.2.24 Theorem. *If M is a closed subspace of a uniformly rotund normed space X, then X/M is uniformly rotund.*

PROOF. This is an immediate consequence of Lemma 5.2.22 and the fact that moduli of rotundity are nondecreasing functions on $[0, 2]$.

Here is another proof that is related to the argument of Lemma 5.2.22 but that does not ultimately rely on Theorem 5.2.5. Suppose that (x_n+M) and $(y_n + M)$ are sequences in $S_{X/M}$ such that $\big\| \frac{1}{2}(x_n + y_n) + M \big\| \to 1$. It suffices to show that $\|(x_n - y_n) + M\| \to 0$. By Proposition 1.7.6, it may be assumed that

$$\|x_n + M\| \le \|x_n\| < \|x_n + M\| + n^{-1}$$

for each n, from which it follows that $\|x_n\| \to 1$. Similarly, it may be assumed that $\|y_n\| \to 1$. Since

$$\big\| \tfrac{1}{2}(x_n + y_n) + M \big\| \le \big\| \tfrac{1}{2}(x_n + y_n) \big\| \le \tfrac{1}{2}\big(\|x_n\| + \|y_n\|\big) \to 1,$$

it also follows that $\big\| \frac{1}{2}(x_n + y_n) \big\| \to 1$, which implies that $\|x_n - y_n\| \to 0$ and therefore that $\|(x_n - y_n) + M\| \to 0$. ■

As is the case for rotundity, uniform rotundity is preserved by the operation of taking direct sums. The proof of this to be given here will seem suspiciously similar to that of Theorem 5.2.11. The reason for this is that both proofs are generalizations of an argument that can be used to derive the uniform rotundity of Euclidean n-space for each positive integer n from the uniform rotundity of the scalar field.

5.2.25 Theorem. *Suppose that $X_1, \dots, X_n$ are normed spaces. Then $X_1 \oplus \cdots \oplus X_n$ is uniformly rotund if and only if each X_j is uniformly rotund.*

PROOF. Suppose first that $X_1 \oplus \cdots \oplus X_n$ is uniformly rotund. Since each X_j is isometrically isomorphic to a subspace of $X_1 \oplus \cdots \oplus X_n$, it follows immediately that each X_j is uniformly rotund.

Now suppose instead that each X_j is uniformly rotund. Let $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ be elements of $S_{X_1 \oplus \cdots \oplus X_n}$ such that

$$\|(x_1, \dots, x_n) - (y_1, \dots, y_n)\| \geq \epsilon > 0.$$

Let

$$A = \left\{ j : j \in \{1, \dots, n\},\ \|x_j - y_j\|^2 \geq \frac{\epsilon^2}{4}(\|x_j\|^2 + \|y_j\|^2) \right\},$$

and observe that $\|x_j - y_j\| \geq (\epsilon/2)\max\{\|x_j\|, \|y_j\|\}$ for each j in A. Let $\lambda(s) = \min\{\delta_{X_1}(s), \dots, \delta_{X_n}(s)\}$ when $0 < s \leq 2$, where $\delta_{X_1}, \dots, \delta_{X_n}$ are the respective moduli of rotundity of the spaces $X_1, \dots, X_n$. With $\gamma_{2,\lambda}$ as in Lemma 5.2.9, it follows from that lemma that

$$\left\|\tfrac{1}{2}(x_j + y_j)\right\|^2 \leq \left(1 - \gamma_{2,\lambda}\left(\frac{\epsilon}{2}\right)\right)\left(\frac{\|x_j\|^2 + \|y_j\|^2}{2}\right)$$

whenever $j \in A$ and that

$$\left\|\tfrac{1}{2}(x_j + y_j)\right\|^2 \leq \frac{\|x_j\|^2 + \|y_j\|^2}{2}$$

whenever $j \in \{1, \dots, n\}$, so

$$\begin{aligned}
1 - &\left\|\tfrac{1}{2}((x_1, \dots, x_n) + (y_1, \dots, y_n))\right\|^2 \\
&= \sum_{j=1}^{n}\left(\frac{\|x_j\|^2 + \|y_j\|^2}{2} - \left\|\tfrac{1}{2}(x_j + y_j)\right\|^2\right) \\
&\geq \sum_{j \in A}\left(\frac{\|x_j\|^2 + \|y_j\|^2}{2} - \left\|\tfrac{1}{2}(x_j + y_j)\right\|^2\right) \\
&\geq \gamma_{2,\lambda}\left(\frac{\epsilon}{2}\right)\sum_{j \in A}\frac{\|x_j\|^2 + \|y_j\|^2}{2}.
\end{aligned}$$

Furthermore, if $B = \{1, \dots, n\} \setminus A$, then

$$\begin{aligned}\sum_{j\in A} \|x_j - y_j\|^2 &= \|(x_1, \dots, x_n) - (y_1, \dots, y_n)\|^2 - \sum_{j\in B} \|x_j - y_j\|^2 \\ &\geq \epsilon^2 - \frac{\epsilon^2}{4} \sum_{j\in B} (\|x_j\|^2 + \|y_j\|^2) \\ &\geq \epsilon^2 - \frac{\epsilon^2}{4} (\|(x_1, \dots, x_n)\|^2 + \|(y_1, \dots, y_n)\|^2) \\ &= \frac{\epsilon^2}{2},\end{aligned}$$

so $(\sum_{j\in A} \|x_j - y_j\|^2)^{1/2} \geq \epsilon/2^{1/2}$, from which it follows that

$$\max\left\{ \left(\sum_{j\in A} \|x_j\|^2\right)^{1/2}, \left(\sum_{j\in A} \|y_j\|^2\right)^{1/2} \right\} \geq \frac{\epsilon}{2 \cdot 2^{1/2}}.$$

Therefore

$$\begin{aligned}1 - \left\|\tfrac{1}{2}((x_1, \dots, x_n) + (y_1, \dots, y_n))\right\|^2 &\geq \gamma_{2,\lambda}\left(\frac{\epsilon}{2}\right) \sum_{j\in A} \frac{\|x_j\|^2 + \|y_j\|^2}{2} \\ &\geq \gamma_{2,\lambda}\left(\frac{\epsilon}{2}\right) \frac{\epsilon^2}{16},\end{aligned}$$

and so

$$\left\|\tfrac{1}{2}((x_1, \dots, x_n) + (y_1, \dots, y_n))\right\| \leq \left(1 - \gamma_{2,\lambda}\left(\frac{\epsilon}{2}\right) \frac{\epsilon^2}{16}\right)^{1/2} < 1.$$

It follows that $X_1 \oplus \cdots \oplus X_n$ is uniformly rotund. ∎

There are many sources available for the reader interested in learning more about uniform rotundity, including [21], [56], [107], [141], and [157], as well as Mahlon Day's papers [47]–[50].

Exercises

5.14 Prove that if X is a normed space and δ_X is its modulus of rotundity, then X is rotund if and only if $\delta_X(2) = 1$.

5.15 Find an explicit formula, in terms only of functions one might encounter in a precalculus course, for the modulus of rotundity δ_2 of Euclidean 2-space. (This formula is important because it provides a simple expression for the least upper bound of the moduli of rotundity of all uniformly rotund normed spaces of dimension at least two. It is a 1960 result of G. Nordlander [176] that if X is a uniformly rotund normed space that is not zero-dimensional if $\mathbb{F} = \mathbb{C}$ and is neither zero- nor one-dimensional if $\mathbb{F} = \mathbb{R}$, and if δ_X is the modulus of rotundity of X, then $\delta_X(\epsilon) \leq \delta_2(\epsilon)$ when $0 \leq \epsilon \leq 2$.)

5.16 This exercise assumes some knowledge of inner product spaces and Hilbert spaces. See, for example, the references cited in Exercise 5.4.

(a) Suppose that X is an inner product space that is not zero-dimensional if $\mathbb{F} = \mathbb{C}$ and is neither zero- nor one-dimensional if $\mathbb{F} = \mathbb{R}$. Find an explicit formula for the modulus of rotundity of X in terms only of functions one might encounter in a precalculus course. (Compare Exercise 5.15.)

(b) Use the formula found in (a) to show that every inner product space is uniformly rotund.

(c) Conclude from (b) that every Hilbert space is reflexive.

5.17 The purpose of this exercise is to provide another proof of Theorem 5.2.11 when $p \geq 2$ that is perhaps a bit more in the spirit of Clarkson's. Let p be such that $2 \leq p < \infty$.

(a) Prove that if $\alpha, \beta \in \mathbb{F}$, then

$$\big(|\alpha+\beta|^p + |\alpha-\beta|^p\big)^{2/p} \leq 2\big(|\alpha|^2 + |\beta|^2\big) \leq 2\big(|\alpha|^p + |\beta|^p\big)^{2/p} 2^{(p-2)/p},$$

and therefore

$$|\alpha+\beta|^p + |\alpha-\beta|^p \leq 2^{p-1}\big(|\alpha|^p + |\beta|^p\big).$$

Notice that this is a generalization of the well-known *parallelogram law* for $\mathbb{F}$, namely, that

$$|\alpha+\beta|^2 + |\alpha-\beta|^2 = 2\big(|\alpha|^2 + |\beta|^2\big)$$

whenever $\alpha, \beta \in \mathbb{F}$.

(b) Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω. Use (a) to prove that $L_p(\Omega, \Sigma, \mu)$ is uniformly rotund.

5.18 This exercise requires James's theorem from either of the optional Sections 1.13 and 2.9. Suppose that X is a uniformly rotund Banach space.

(a) Show that for each nonzero member x^* of X^*, the set

$$\{\, x : x \in X,\ \mathrm{Re}\, x^*x = \|x^*\| \,\}$$

has a unique point closest to the origin. Do not use the Milman-Pettis theorem or its corollaries to do this (but do look at the behavior of the nets in the proof of the Milman-Pettis theorem).

(b) Derive the Milman-Pettis theorem from (a) and James's theorem.

5.19 Suppose that the requirement that M be a set of existence is added to the hypotheses of Lemma 5.2.22. Then there is a much simpler proof of the lemma that does not rely on Theorem 5.2.5 and that extends to the case in which $\epsilon = 2$. Find this proof.

It can be shown that every reflexive subspace of a normed space is a set of existence; see Exercise 5.12. Since every closed subspace of a reflexive normed space is reflexive, it follows that Lemma 5.2.22 can be extended to the case in which $\epsilon = 2$ under the additional hypothesis that X is reflexive.

5.20 (M. M. Day, 1944 [50]). Prove that a normed space X is uniformly rotund if and only if

$$\inf\{\,\delta_M(\epsilon) : M \text{ is a two-dimensional subspace of } X\,\} > 0$$

whenever $0 < \epsilon \le 2$, where δ_M is the modulus of rotundity of the subspace M.

5.21 (M. M. Day, 1941 [48]). This exercise uses the notion of the ℓ_p sum of normed spaces developed in Exercise 5.1. Suppose that (X_n) is a sequence of normed spaces, that (δ_{X_n}) is the corresponding sequence of moduli of rotundity of the spaces, and that $1 < p < \infty$. Prove that $\ell_p((X_n))$ is uniformly rotund if and only if each X_n is uniformly rotund and $\inf_n \delta_{X_n}(\epsilon) > 0$ when $0 < \epsilon \le 2$. (The condition that each X_n be uniformly rotund has been added only for emphasis, since the condition involving the infimum of the moduli of rotundity clearly implies it.)

5.22 Suppose that X is a superreflexive Banach space and that Y is a Banach space such that there is a bounded linear operator from X onto Y. Prove that Y is superreflexive.

5.23 Prove that the Banach space X of Example 5.2.13 is superreflexive. (Exercise 5.21 might be helpful.) Conclude that X is a reflexive Banach space that is rotund but not uniformly rotund.

5.3 Generalizations of Uniform Rotundity

In approximation theory, it is sometimes important to know if a Banach space X has the property that whenever x is an element of X and (x_n) is a sequence in a closed convex subset C of X such that $\|x - x_n\| \to d(x, C)$, then (x_n) converges. For the moment, call this property D; the definition will appear again later in an equivalent form. As will be seen below, for a Banach space X to have property D it is *necessary* that X be rotund and *sufficient* that X be uniformly rotund. However, it is known that there are rotund Banach spaces lacking property D and Banach spaces that have property D without being uniformly rotund. Thus, property D can be viewed as a strong form of rotundity that lies properly between simple rotundity and uniform rotundity.

Many other properties lying properly between rotundity and uniform rotundity have been defined and studied since Clarkson's introduction of the notions of rotundity and uniform rotundity in 1936. The purpose of this section is to examine a few of the more well-known ones. Since such a study tends to abound with lengthy terms such as "weakly locally uniformly rotund" and "uniformly rotund in every direction," it is both helpful and customary to use abbreviations for these terms along with the following notational scheme.

5.3.1 Notation. Suppose that P is an abbreviation for a property that normed spaces can have. Then a normed space X is said to be (P) if it has property P, and is said to be $\langle$P$\rangle$ if there is a norm $\|\cdot\|_{\mathrm{P}}$ for X equivalent to its original norm such that $(X, \|\cdot\|_{\mathrm{P}})$ is (P). In particular, a normed space is said to be (UR), (R), (B), (Rf), or (H) if it is respectively uniformly rotund, rotund, a Banach space, reflexive, or a Radon-Riesz space. A normed space that has properties P and Q is said to be (P) & (Q).

For example, the ℓ_1^2 and ℓ_2^2 norms are equivalent on $\mathbb{F}^2$, so ℓ_1^2 is not (R) but is $\langle$UR$\rangle$. As further examples, the statement that every uniformly rotund Banach space is a rotund reflexive Radon-Riesz space can be abbreviated to (UR) & (B) $\Rightarrow$ (R) & (Rf) & (H), while the fact that every superreflexive Banach space is reflexive can be stated symbolically as $\langle$UR$\rangle$ & (B) $\Rightarrow$ (Rf).

The first generalization of uniform rotundity to be examined here is a localization of that property obtained by requiring that for each fixed x in the unit sphere of a normed space X and each positive ϵ, there is a positive δ depending on ϵ *and* x such that $\left\|\frac{1}{2}(x+y)\right\| \leq 1-\delta$ whenever $y \in S_X$ and $\|x-y\| \geq \epsilon$.

5.3.2 Definition. (A. R. Lovaglia, 1955 [159]). Suppose that X is a normed space. Define a function $\delta_X \colon [0,2] \times S_X \to [0,1]$ by the formula

$$\delta_X(\epsilon, x) = \inf\left\{\, 1 - \left\|\tfrac{1}{2}(x+y)\right\| : y \in S_X,\ \|x-y\| \geq \epsilon \,\right\}.$$

Then δ_X is the *LUR modulus* of X. The space X is *locally uniformly rotund* or *locally uniformly convex* if $\delta_X(\epsilon, x) > 0$ whenever $0 < \epsilon \leq 2$ and $x \in S_X$. The abbreviation LUR is used for this property.

Notice that every zero-dimensional normed space is (LUR) since the requirement on its LUR modulus is satisfied vacuously. Similar observations about the other moduli defined later in this section show that zero-dimensional normed spaces have the rotundity properties defined in terms of those moduli.

A comparison of the definitions of uniform rotundity and local uniform rotundity shows that the first implies the second. Also, if X is a locally uniformly rotund normed space with LUR modulus δ_X and x and y are different elements of S_X, then $\left\|\frac{1}{2}(x+y)\right\| \leq 1 - \delta_X\left(\|x-y\|, x\right) < 1$, which finishes the proof of the following result.

5.3.3 Proposition. *Every uniformly rotund normed space is locally uniformly rotund, and every locally uniformly rotund normed space is rotund. In symbols,* (UR) $\Rightarrow$ (LUR) $\Rightarrow$ (R).

In particular, the uniformly rotund normed space ℓ_2^2 is locally uniformly rotund though it is isomorphic to the nonrotund normed space ℓ_1^2, so local uniform rotundity is not preserved by isomorphisms. It certainly is by isometric isomorphisms.

5.3.4 Proposition. *Every normed space that is isometrically isomorphic to a locally uniformly rotund normed space is itself locally uniformly rotund.*

In general, the strengthenings of rotundity to be studied in this section are preserved by isometric isomorphisms but not isomorphisms, though this will not be stated explicitly for any more of these properties. The argument for isometric isomorphisms will always be obvious, while the argument for isomorphisms can always be based on ℓ_2^2 as above.

As with uniform rotundity, there are sequential characterizations of local uniform rotundity that are often used as its definition. Notice that the proof of the following result is essentially the same as that of Proposition 5.2.8.

5.3.5 Proposition. *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is locally uniformly rotund.*
(b) *When $x \in S_X$ and (y_n) is a sequence in S_X such that $\left\|\frac{1}{2}(x+y_n)\right\| \to 1$, it follows that $\|x-y_n\| \to 0$.*
(c) *When $x \in S_X$ and (y_n) is a sequence in B_X such that $\left\|\frac{1}{2}(x+y_n)\right\| \to 1$, it follows that $\|x-y_n\| \to 0$.*
(d) *When $x \in S_X$ and (y_n) is a sequence in X such that $\|y_n\|$ and $\left\|\frac{1}{2}(x+y_n)\right\|$ both tend to 1, it follows that $\|x-y_n\| \to 0$.*

PROOF. Suppose that (b) holds and that x is an element of S_X and (y_n) a sequence in X such that $\|y_n\|$ and $\left\|\frac{1}{2}(x+y_n)\right\|$ both tend to 1. It will be shown that $\|x-y_n\| \to 0$. By discarding terms from the beginning of the sequence if necessary, it may be assumed that no y_n is 0. Then

$$1 \geq \left\|\tfrac{1}{2}\left(x+\|y_n\|^{-1}y_n\right)\right\| \geq \left\|\tfrac{1}{2}(x+y_n)\right\| - \left\|\tfrac{1}{2}(1-\|y_n\|^{-1})y_n\right\| \to 1,$$

so $\left\|\frac{1}{2}\left(x+\|y_n\|^{-1}y_n\right)\right\| \to 1$. Since $\|y_n\|^{-1}y_n \to x$ by (b), it follows that $y_n \to x$, which establishes that (b) $\Rightarrow$ (d).

Now suppose that (d) holds and that x is an element of S_X and (y_n) a sequence in B_X such that $\left\|\frac{1}{2}(x+y_n)\right\| \to 1$. Since

$$\left\|\tfrac{1}{2}(x+y_n)\right\| \leq \tfrac{1}{2}\left(1+\|y_n\|\right) \leq 1$$

for each n, it follows that $\|y_n\| \to 1$, so $\|x-y_n\| \to 0$ by (d). Therefore (d) $\Rightarrow$ (c), from which it immediately follows that (b) $\Leftrightarrow$ (c) $\Leftrightarrow$ (d).

Suppose that X is locally uniformly rotund and that x is an element of S_X and (y_n) a sequence in S_X such that $\left\|\frac{1}{2}(x+y_n)\right\| \to 1$ but $\|x-y_n\|$ does not tend to 0. Let δ_X be the LUR modulus of X. It follows that there is a subsequence (y_{n_j}) of (y_n) such that $\|x-y_{n_j}\| \geq \epsilon$ for some positive ϵ and each j, which implies that $\left\|\frac{1}{2}(x+y_{n_j})\right\| \leq 1-\delta_X(\epsilon,x)$ for each j, a contradiction. Therefore (a) $\Rightarrow$ (b).

Finally, suppose that X is not locally uniformly rotund. Then there is an ϵ such that $0 < \epsilon \le 2$ and an x in S_X for which $\delta_X(\epsilon, x) = 0$. Therefore there is a sequence (y_n) in S_X such that $\|x - y_n\| \ge \epsilon$ for each n but $\left\|\frac{1}{2}(x + y_n)\right\| \to 1$, so (b) does not hold. This shows that (b) $\Rightarrow$ (a). ■

For most of the generalizations of uniform rotundity in this section, examples will be cited from the work of Mark Smith to show that the properties are distinct from each other and from rotundity and uniform rotundity. Space limitations prevent the detailed presentation of all the examples, but the following one showing that local uniform rotundity does not imply uniform rotundity will give the flavor of such constructions.

5.3.6 Example. (M. A. Smith, 1978 [219]). This example uses the *parallelogram law* for ℓ_2, namely, that if $(\alpha_n), (\beta_n) \in \ell_2$, then

$$\|(\alpha_n) + (\beta_n)\|_2^2 + \|(\alpha_n) - (\beta_n)\|_2^2 = 2\left(\|(\alpha_n)\|_2^2 + \|(\beta_n)\|_2^2\right).$$

See Exercise 5.24. Also needed is the fact that if $(\alpha_n) \in \ell_1$, then $(\alpha_n) \in \ell_2$ and $\|(\alpha_n)\|_2 \le \|(\alpha_n)\|_1$. The proof of this is easy: If $(\alpha_n) \in S_{\ell_1}$, then $|\alpha_n| \le 1$ for each n, so $\sum_n |\alpha_n|^2 \le \sum_n |\alpha_n| = 1$ and $(\alpha_n) \in B_{\ell_2}$.

For each member (α_n) of ℓ_1, let

$$\|(\alpha_n)\|_E = \left(\|(\alpha_n)\|_1^2 + \|(\alpha_n)\|_2^2\right)^{1/2}.$$

It is easy to check that $\|\cdot\|_E$ is a norm on ℓ_1, and this norm is equivalent to the usual norm of ℓ_1 since $\|(\alpha_n)\|_1 \le \|(\alpha_n)\|_E \le \sqrt{2}\,\|(\alpha_n)\|_1$ whenever $(\alpha_n) \in \ell_1$.

Suppose that $x \in S_{(\ell_1, \|\cdot\|_E)}$ and that (y_k) is a sequence of elements of $S_{(\ell_1, \|\cdot\|_E)}$ such that $\left\|\frac{1}{2}(x + y_k)\right\|_E \to 1$. For notational convenience, the element x and each y_k will sometimes be denoted by (α_n) and $(\beta_{n,k})$ respectively. If u and v are members of a normed space, then

$$\begin{aligned}\|u + v\|^2 &\le \left(\|u\| + \|v\|\right)^2 \\ &\le \left(\|u\| + \|v\|\right)^2 + \left(\|u\| - \|v\|\right)^2 \\ &= 2\left(\|u\|^2 + \|v\|^2\right),\end{aligned}$$

so $2\left(\|u\|^2 + \|v\|^2\right) - \|u + v\|^2 \ge 0$. Since

$$\begin{aligned}0 &\le 2\left(\|x\|_1^2 + \|y_k\|_1^2\right) - \|x + y_k\|_1^2 + 2\left(\|x\|_2^2 + \|y_k\|_2^2\right) - \|x + y_k\|_2^2 \\ &= 2\left(\|x\|_E^2 + \|y_k\|_E^2\right) - \|x + y_k\|_E^2 \\ &= 4 - \|x + y_k\|_E^2\end{aligned}$$

for each k, and since $4 - \|x + y_k\|_E^2 \to 0$, it follows that

$$2\left(\|x\|_j^2 + \|y_k\|_j^2\right) - \|x + y_k\|_j^2 \to 0$$

when $j = 1, 2$. This, together with the fact that

$$\begin{aligned} 0 &\le \left(\|x\|_1 - \|y_k\|_1\right)^2 \\ &= 2\left(\|x\|_1^2 + \|y_k\|_1^2\right) - \left(\|x\|_1 + \|y_k\|_1\right)^2 \\ &\le 2\left(\|x\|_1^2 + \|y_k\|_1^2\right) - \|x + y_k\|_1^2 \end{aligned}$$

for each k, shows that $\|y_k\|_1 \to \|x\|_1$. Moreover, since

$$\|x - y_k\|_2^2 = 2\left(\|x\|_2^2 + \|y_k\|_2^2\right) - \|x + y_k\|_2^2$$

for each k by the parallelogram law, it also follows that $\|x - y_k\|_2 \to 0$, so $\beta_{n,k} \to \alpha_n$ for each n.

It will now be shown that $\|x - y_k\|_1 \to 0$. If $x = 0$, then this follows immediately from the fact that $\|y_k\|_1 \to \|x\|_1$. Suppose instead that $x \neq 0$. Fix a positive ϵ less than $\|x\|_1$ and let m be a positive integer such that $\sum_{n=m+1}^{\infty} |\alpha_n| < \epsilon$. Notice that $\sum_{n=1}^{m} |\alpha_n| > \|x\|_1 - \epsilon$. It follows that $\sum_{n=1}^{m} |\beta_{n,k}| > \|x\|_1 - \epsilon$ for large k, which together with the fact that $\sum_n |\beta_{n,k}| = \|y_k\|_1 \to \|x\|_1$ implies that $\sum_{n=m+1}^{\infty} |\beta_{n,k}| < 2\epsilon$ for large k. Therefore for large k,

$$\|x - y_k\|_1 \le \sum_{n=1}^{m} |\alpha_n - \beta_{n,k}| + \sum_{n=m+1}^{\infty} |\alpha_n| + \sum_{n=m+1}^{\infty} |\beta_{n,k}| < \epsilon + \epsilon + 2\epsilon = 4\epsilon,$$

which shows that $\|x - y_k\|_1 \to 0$.

It follows from the equivalence of $\|\cdot\|_1$ and $\|\cdot\|_E$ that $\|x - y_k\|_E \to 0$, which establishes that $(\ell_1, \|\cdot\|_E)$ is locally uniformly rotund. By the Milman-Pettis theorem, this Banach space is not uniformly rotund, or even isomorphic to a uniformly rotund normed space, since it is not reflexive. The space does have the Radon-Riesz property, and even Schur's property, since $(\ell_1, \|\cdot\|_1)$ has Schur's property, so $(\ell_1, \|\cdot\|_E)$ is an example of a nonreflexive Banach space that is not ⟨UR⟩ but is (LUR) & (H).

Smith's paper [219] also contains examples of reflexive Banach spaces that are locally uniformly rotund without being uniformly rotund, as well as reflexive and nonreflexive Banach spaces that are rotund but not locally uniformly rotund.

Although an issue was made of the fact that the Banach space of the preceding example has the Radon-Riesz property, this is actually true for every locally uniformly rotund normed space. Notice that the statement and proof of the following result are obtained from the statement and proof of Theorem 5.2.18 by just substituting "locally uniformly rotund" for "uniformly rotund" everywhere it occurs.

5.3.7 Theorem. (R. Vyborny, 1956 [240]). *Every locally uniformly rotund normed space has the Radon-Riesz property. In symbols,* (LUR) $\Rightarrow$ (H).

PROOF. Suppose that X is a locally uniformly rotund normed space and that (x_n) is a sequence in X and x an element of X such that $x_n \xrightarrow{w} x$ and $\|x_n\| \to \|x\|$. It is to be shown that $x_n \to x$, so it may be assumed that $x \neq 0$ and therefore, after discarding some initial terms of (x_n) if necessary, that each x_n is nonzero. Now $\|x_n\|^{-1}x_n \xrightarrow{w} \|x\|^{-1}x$, and it is enough to show that $\|x_n\|^{-1}x_n \to \|x\|^{-1}x$, so it may even be assumed that x and each x_n are in S_X. Since $\frac{1}{2}(x_n+x) \xrightarrow{w} x$, the weak lower semicontinuity of the norm function implies that $\|\frac{1}{2}(x_n+x)\| \to 1$, so $\|x_n - x\| \to 0$, as required. ■

One reason for the importance of local uniform rotundity is that it imparts some of the same benefits, such as the presence of the Radon-Riesz property, as does uniform rotundity, while far more spaces can be equivalently renormed to be (LUR) than to be (UR). It was shown in a 1971 paper by S. L. Troyanski [232] that every weakly compactly generated Banach space is ⟨LUR⟩, while no nonreflexive Banach space can be ⟨UR⟩.

The next generalization of uniform rotundity is obtained by letting the weak topology play the role of the norm topology in the definition of uniform rotundity, in a sense that will be made clearer by Proposition 5.3.9.

5.3.8 Definition. (V. L. Šmulian, 1939, 1940 [223, 224]). Suppose that X is a normed space. Define a function $\delta_X\colon [0,2] \times S_{X^*} \to [0,1]$ by the formula

$$\delta_X(\epsilon, x^*) = \inf\left(\{1\} \cup \left\{\, 1 - \left\|\tfrac{1}{2}(x+y)\right\| : x, y \in S_X,\ |x^*(x-y)| \geq \epsilon \,\right\}\right).$$

Then δ_X is the *wUR modulus* of X. The space X is *weakly uniformly rotund* or *weakly uniformly convex* if $\delta_X(\epsilon, x^*) > 0$ whenever $0 < \epsilon \leq 2$ and $x^* \in S_{X^*}$. The abbreviation wUR is used for this property.

The reason for specifically including 1 in the set whose infimum defines the wUR modulus is to keep the modulus finite-valued in one particular situation. With all notation as in Definition 5.3.8, suppose that some x^* in S_{X^*} is not norm-attaining. Then there are no members x and y of S_X such that $|x^*(x-y)| \geq 2$, so $\delta_X(2, x^*)$ would be $+\infty$ were it not for the 1.

As with uniform rotundity and local uniform rotundity, weak uniform rotundity has sequential characterizations. The next result gives one along the lines of the equivalence of (a) and (b) in Propositions 5.2.8 and 5.3.5. The derivation of other characterizations analogous to parts (c) and (d) of those two propositions is left as an exercise for the interested reader.

5.3.9 Proposition. *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is weakly uniformly rotund.*
(b) *Whenever (x_n) and (y_n) are sequences in S_X and $\|\frac{1}{2}(x_n+y_n)\| \to 1$, it follows that $x_n - y_n \xrightarrow{w} 0$.*

PROOF. Suppose that (b) fails. Then there are sequences (x_n) and (y_n) in S_X and an x^* in S_{X^*} such that $\left\|\frac{1}{2}(x_n + y_n)\right\| \to 1$ but $x^*(x_n - y_n)$ does not tend to 0. Let δ_X be the wUR modulus of X. It follows that there is a subsequence (x_{n_j}) of (x_n) such that $|x^*(x_{n_j} - y_{n_j})| \geq \epsilon$ for some positive ϵ and each j, implying that $\delta_X(\epsilon, x^*) \leq 1 - \left\|\frac{1}{2}(x_{n_j} + y_{n_j})\right\|$ for each j and therefore that $\delta_X(\epsilon, x^*) = 0$. Therefore X is not weakly uniformly rotund.

Suppose conversely that X is not weakly uniformly rotund. Then there is an ϵ such that $0 < \epsilon \leq 2$ and an x^* in S_{X^*} for which $\delta_X(\epsilon) = 0$. Therefore there are sequences (x_n) and (y_n) in S_X such that $|x^*(x_n - y_n)| \geq \epsilon$ for each n but $\left\|\frac{1}{2}(x_n + y_n)\right\| \to 1$, so (b) does not hold. ■

It follows easily from the preceding proposition that weak uniform rotundity lies between uniform rotundity and rotundity.

5.3.10 Proposition. *Every uniformly rotund normed space is weakly uniformly rotund, and every weakly uniformly rotund normed space is rotund. In symbols,* (UR) $\Rightarrow$ (wUR) $\Rightarrow$ (R).

PROOF. Suppose that X is a uniformly rotund normed space. If (x_n) and (y_n) are sequences in S_X such that $\left\|\frac{1}{2}(x_n + y_n)\right\| \to 1$, then $x_n - y_n \to 0$ by Proposition 5.2.8, so $x_n - y_n \xrightarrow{w} 0$. It follows from Proposition 5.3.9 that X is weakly uniformly rotund.

Now suppose that X is a normed space that is not rotund. Then there are distinct elements x and y of X such that $\|x\| = \|y\| = \left\|\frac{1}{2}(x + y)\right\| = 1$. Let x^* be a member of X^* such that $x^*(x - y) \neq 0$ and let $x_n = x$ and $y_n = y$ for each positive integer n. Then $x^*(x_n - y_n)$ does not tend to 0, which implies that $x_n - y_n$ does not tend weakly to 0. An application of Proposition 5.3.9 shows that X is not weakly uniformly rotund. ■

Mark Smith's paper [219] contains examples of Banach spaces $(\ell_2, \|\cdot\|_L)$ and $(\ell_2, \|\cdot\|_W)$, each formed by putting a norm on ℓ_2 equivalent to its original norm, such that $(\ell_2, \|\cdot\|_L)$ is (LUR) but not (wUR) while $(\ell_2, \|\cdot\|_W)$ is (wUR) without being (LUR). This shows that neither of the conditions (LUR) and (wUR) implies the other, and also shows that weak uniform rotundity does not imply uniform rotundity. Smith also shows that $(\ell_2, \|\cdot\|_W)$ lacks the Radon-Riesz property, so (wUR) does not imply (H).

Incidentally, it turns out that the Banach space $(\ell_1, \|\cdot\|_E)$ of Example 5.3.6 is another Banach space that is (LUR) but not (wUR), since V. E. Zizler showed in a 1971 paper [248] that ℓ_1 cannot be equivalently renormed to be (wUR).

There is an obvious analog of weak uniform rotundity for dual spaces that is obtained by exchanging the roles of the normed space and its dual in the definition of the wUR modulus. This analog will be examined only briefly here, but will have an important application in Section 5.6.

5.3.11 Definition. (V. L. Šmulian, 1939, 1940 [223, 224]). Suppose that X is a normed space. Define a function $\delta_{X^*}\colon [0,2] \times S_X \to [0,1]$ by the formula

$$\delta_{X^*}(\epsilon, x) = \inf\left\{\, 1 - \left\|\tfrac{1}{2}(x^* + y^*)\right\| : x^*, y^* \in S_{X^*},\ |(x^* - y^*)(x)| \geq \epsilon \,\right\}.$$

Then δ_{X^*} is the *w*UR modulus* of X^*. The space X^* is *weakly* uniformly rotund* or *weakly* uniformly convex* if $\delta_{X^*}(\epsilon, x) > 0$ whenever $0 < \epsilon \leq 2$ and $x \in S_X$. The abbreviation w*UR is used for this property.

Notice that unlike what was done for the wUR modulus, there is no 1 explicitly included in the set whose infimum is being taken to obtain the w*UR modulus. This is not needed to assure that $\delta_{X^*}(2, x)$ is finite when $x \in S_X$, since for each x in S_X there is an x^* in S_{X^*} such that $x^*x = 1$ and therefore such that $\left|(x^* - (-x^*))(x)\right| = 2$.

It should be noted that some sources say that if X is a normed space such that X^* satisfies the above definition of weak* uniform rotundity, then it is X instead of X^* that is called weakly* uniformly rotund. See [107, p. 71] for an instance of the use of the term in this alternative sense, and [239, p. 48] for an example of the use of the term the way it has been defined here.

5.3.12 Proposition. *Suppose that X is a normed space. If X^* is weakly uniformly rotund, then it is weakly* uniformly rotund, and if X^* is weakly* uniformly rotund, then it is rotund. In symbols,* (wUR) $\Rightarrow$ (w*UR) $\Rightarrow$ (R) *for dual spaces of normed spaces.*

PROOF. Let Q be the natural map from X into X^{**}. Then the w*UR modulus of X^* is given by the formula

$$\delta_{X^*}(\epsilon, x) = \inf\left\{\, 1 - \left\|\tfrac{1}{2}(x^* + y^*)\right\| : x^*, y^* \in S_{X^*},\ |(Qx)(x^* - y^*)| \geq \epsilon \,\right\}$$

whenever $0 \leq \epsilon \leq 2$ and $x \in S_X$. Comparing this to the formula for the wUR modulus of X^* shows that X^* is weakly* uniformly rotund if it is weakly uniformly rotund.

Now suppose that X^* is not rotund. Let x_1^* and x_2^* be different elements of S_{X^*} such that $\frac{1}{2}(x_1^* + x_2^*) \in S_{X^*}$ and let x_0 be an element of S_X such that $x_1^*x_0 \neq x_2^*x_0$. Let $\epsilon_0 = |(x_1^* - x_2^*)(x_0)|$. Then $\delta_{X^*}(\epsilon_0, x_0) = 0$, so X^* is not weakly* uniformly rotund. ■

Local uniform rotundity and weak uniform rotundity have a common generalization stronger than simple rotundity. Given the definitions of local and weak uniform rotundity, the definition of this new property could be deduced from its name.

5.3.13 Definition. (A. R. Lovaglia, 1955 [159]). Suppose that X is a normed space. Define a function $\delta_X\colon [0,2] \times S_X \times S_{X^*} \to [0,1]$ by the formula

$$\delta_X(\epsilon, x, x^*) = \inf\left(\{1\} \cup \left\{ 1 - \left\|\tfrac{1}{2}(x+y)\right\| : y \in S_X,\ |x^*(x-y)| \geq \epsilon \right\}\right).$$

Then δ_X is the *wLUR modulus* of X. The space X is *weakly locally uniformly rotund* or *weakly locally uniformly convex* if $\delta_X(\epsilon, x, x^*) > 0$ whenever $0 < \epsilon \leq 2$, $x \in S_X$, and $x^* \in S_{X^*}$. The abbreviation wLUR is used for this property.

A glance at the definitions of local, weak, and weak local uniform rotundity shows that each of the first two properties implies the third. Also, if X is a normed space with wLUR modulus δ_X and there are distinct elements x and y of X such that $\|x\| = \|y\| = \left\|\frac{1}{2}(x+y)\right\| = 1$, then there is an x^* in S_{X^*} such that $x^*(x-y) \neq 0$, which implies that $\delta_X\left(|x^*(x-y)|, x, x^*\right) = 0$. It follows that weak local uniform rotundity implies rotundity. These observations are summarized in the following proposition.

5.3.14 Proposition. *Every normed space that is either locally uniformly rotund or weakly uniformly rotund is weakly locally uniformly rotund, and every weakly locally uniformly rotund normed space is rotund. In symbols,* (LUR) $\Rightarrow$ (wLUR) *and* (wUR) $\Rightarrow$ (wLUR) $\Rightarrow$ (R).

Mark Smith's paper [219] has examples showing that none of the implications in the preceding proposition is reversible.

The next generalization of uniform rotundity is defined in terms of geometric properties of convex sets instead of the behavior of a modulus. Recall that the *diameter* of a nonempty subset A of a metric space is given by the formula $\operatorname{diam}(A) = \sup\{\, d(x,y) : x, y \in A \,\}$.

5.3.15 Definition. (V. L. Šmulian, 1940 [224]; K. Fan and I. Glicksberg, 1958 [77]). Suppose that X is a normed space such that whenever C is a nonempty convex subset of X, the diameter of $C \cap tB_X$ tends to 0 as t decreases to $d(0, C)$. Then X is *strongly rotund* or *strongly convex*. The abbreviation K is used for this property.

The abbreviation D is sometimes used for the combination of strong rotundity and completeness; that is, (D) $\Leftrightarrow$ (K) & (B).

5.3.16 Proposition. *Every uniformly rotund normed space is strongly rotund, and every strongly rotund normed space is rotund. In symbols,* (UR) $\Rightarrow$ (K) $\Rightarrow$ (R).

PROOF. Suppose first that X is a normed space that is not strongly rotund. Let C be a nonempty convex subset of X such that $\operatorname{diam}(C \cap tB_X)$ does not tend to 0 as t decreases to $d(0, C)$. Clearly $d(0, C) \neq 0$, so it may be assumed that $d(0, C) = 1$. It follows that there must be a positive ϵ such that for each positive integer n there are elements x_n and y_n of $C \cap (1+n^{-1})B_X$ such that $\|x_n - y_n\| \geq \epsilon$. Now $\frac{1}{2}(x_n + y_n) \in C \cap (1+n^{-1})B_X$ for each n by convexity, from which it follows that $\|x_n\|$, $\|y_n\|$, and $\|\frac{1}{2}(x_n + y_n)\|$ all tend to 1 even though $\|x_n - y_n\|$ does not tend to 0. By Proposition 5.2.8, the space X is not uniformly rotund.

Now suppose instead that X is a nonrotund normed space. Then there are distinct elements x and y of S_X such that $\{\, tx + (1-t)y : 0 \leq t \leq 1 \,\}$ lies in S_X. This line segment is a nonempty convex subset L of X such that $\operatorname{diam}(L \cap tB_X)$ does not tend to 0 as t decreases to $d(0, L)$, so X is not strongly rotund. ■

Each of the two characterizations of strong rotundity in the following theorem is sometimes used as its definition. In fact, the condition given in (c) is essentially Šmulian's original formulation of the property.

5.3.17 Theorem. (K. Fan and I. Glicksberg, 1958 [77]). *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is strongly rotund.*

(b) *Whenever $x^* \in S_{X^*}$, the diameter of $\{\, x : x \in B_X, \operatorname{Re} x^*x \geq 1-\delta \,\}$ tends to 0 as δ decreases to 0.*

(c) *Whenever (x_n) is a sequence in S_X for which there is an element x^* of S_{X^*} such that $\operatorname{Re} x^*x_n \to 1$, the sequence (x_n) is Cauchy.*

PROOF. Suppose first that (b) does not hold. This implies the existence of an x^* in S_{X^*} and a positive ϵ such that for each positive integer n there are elements x_n and y_n of B_X for which $\operatorname{Re} x^*x_n \geq (1+n^{-1})^{-1}$, $\operatorname{Re} x^*y_n \geq (1+n^{-1})^{-1}$, and $\|x_n - y_n\| \geq \epsilon$. Let $A = \{\, x : x \in X, \operatorname{Re} x^*x \geq 1 \,\}$, a nonempty convex subset of X. It is clear that $d(0, A) \geq 1$. Since there is a sequence (z_n) of members of B_X such that $\operatorname{Re} x^*z_n > 0$ for each n and $\operatorname{Re} x^*z_n \to 1$, and since $\big((\operatorname{Re} x^*z_n)^{-1}z_n\big)$ is a sequence in A such that $\|(\operatorname{Re} x^*z_n)^{-1}z_n\| \to 1$, it follows that $d(0, A) = 1$. Let $u_n = (1+n^{-1})x_n$ and $v_n = (1+n^{-1})y_n$ for each n. Then $u_n, v_n \in A \cap (1+n^{-1})B_X$ and $\|u_n - v_n\| > \epsilon$ for each positive integer n, so $\operatorname{diam}(A \cap tB_X)$ does not tend to 0 as t decreases to $d(0, A)$. Therefore X is not strongly rotund, which shows that (a) $\Rightarrow$ (b).

It is clear that (b) $\Rightarrow$ (c). To see that (c) $\Rightarrow$ (a), suppose that X is not strongly rotund. Then there is a nonempty convex subset C of X such that $d(0, C) = 1$ and $\operatorname{diam}(C \cap tB_X)$ does not tend to 0 as t decreases to 1. Therefore there must be a positive γ and a sequence (w_n) in C such that $\|w_n\| \to 1$ and $\|w_{2n-1} - w_{2n}\| \geq \gamma$ for each positive integer n. By Eidelheit's separation theorem, there is an x^* in X^* and a positive real number s such that $\operatorname{Re} x^*x \geq s$ whenever $x \in C$ and $\operatorname{Re} x^*x \leq s$ whenever $x \in B_X$. It may be assumed that $s = 1$. Then $\|x^*\| = \|\operatorname{Re} x^*\| \leq 1$, and in fact $\|x^*\| = 1$ since $\operatorname{Re} x^*\big(\|w_n\|^{-1}w_n\big) \to 1$; for this, notice that

$$1 \geq \operatorname{Re} x^*\big(\|w_n\|^{-1}w_n\big) \geq \|w_n\|^{-1}$$

for each n. Now $(\|w_n\|^{-1}w_n)$ is a sequence in S_X that cannot be Cauchy since $\|w_n\| \to 1$ and (w_n) is not Cauchy. Since $\operatorname{Re} x^*(\|w_n\|^{-1}w_n) \to 1$, it follows that (c) does not hold. Therefore (c) $\Rightarrow$ (a). ■

A normed space that satisfies condition (b) of the preceding theorem is sometimes said to be (ν). Day's wonderful description of this property in [56] is that "thin nibbles are small nibbles."

5.3.18 Corollary. *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is a strongly rotund Banach space.*

(b) *Whenever (x_n) is a sequence in S_X for which there is an element x^* of S_{X^*} such that $\operatorname{Re} x^*x_n \to 1$, the sequence (x_n) converges.*

PROOF. It follows immediately from the preceding theorem that (a) $\Rightarrow$ (b) and that (b) implies that X is strongly rotund. All that needs to be shown is that (b) implies that X is complete. Suppose that (b) holds and that (y_n) is a Cauchy sequence in X. It may be assumed that $\|y_n\|$ does not tend to 0 and therefore that $\|y_n\| \to 1$ and no y_n is zero. Since there is a norm-one element of the completion of X to which (y_n) converges, and since there is a norm-one member of the dual space of the completion of X that maps this limit to 1, there is an x^* in S_{X^*} such that $\operatorname{Re} x^*(\|y_n\|^{-1}y_n) \to 1$. Therefore $(\|y_n\|^{-1}y_n)$ has a limit, and so (y_n) has the same limit. ■

In the branch of approximation theory known as nearest point theory in Banach spaces, it is often useful to know when a Banach space is strongly rotund. This is because of several characterizations of the property that use the terms of the following definition.

5.3.19 Definition. Suppose that A is a nonempty subset of a metric space M, that $x \in M$, and that (y_n) is a sequence in A such that $d(x, y_n) \to d(x, A)$. Then (y_n) is a *minimizing sequence* in A with respect to x. The set A is *approximatively compact* if each minimizing sequence in A has a convergent subsequence whose limit is in A.

The word "approximatively" is occasionally misspelled "approximately" in the literature, since that is what the eye tends to see at a rapid glance. While the idea of a minimizing sequence is an old one, the notion of approximative compactness was introduced in a 1961 paper by N. V. Efimov and S. B. Stechkin [72].

One of the reasons for the importance of approximative compactness in approximation theory is that every approximatively compact set A is a set of existence, and therefore for each element x of the space there is at least one point of A that is a best approximation to x from the set in the sense that it is at least as close to x as every other member of A. To see that approximatively compact sets are always sets of existence, suppose that A is an approximatively compact subset of a metric space M and that $x \in M$. Then there is a minimizing sequence (y_n) in A with respect to x; just select each y_n from A so that $d(x, y_n) \leq d(x, A) + n^{-1}$. Then (y_n) has a subsequence that converges to some y in A, and $d(x, y) = d(x, A)$.

5.3.20 Theorem. (K. Fan and I. Glicksberg, 1958 [77]). *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is strongly rotund.*

(b) *Whenever C is a nonempty convex subset of X and (y_n) is a minimizing sequence in C with respect to some x in X, the sequence (y_n) is Cauchy.*

PROOF. It is very easily seen that (b) is equivalent to the following statement.

(b_0) Whenever C is a nonempty convex subset of X and (y_n) is a minimizing sequence in C with respect to 0, the sequence (y_n) is Cauchy.

The equivalence of (a) and (b_0) follows almost immediately from the definition of strong rotundity. ■

A space that satisfies part (b) of the preceding theorem is sometimes said to be (K_ω).

5.3.21 Theorem. (K. Fan and I. Glicksberg, 1958 [77]). *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is a strongly rotund Banach space.*

(b) *Whenever C is a nonempty convex subset of X and (y_n) is a minimizing sequence in C with respect to some x in X, the sequence (y_n) converges.*

(c) *Whenever C is a nonempty closed convex subset of X and (y_n) is a minimizing sequence in C with respect to some x in X, the sequence (y_n) converges to an element y of C.*

(d) *Every nonempty closed convex subset of X is an approximatively compact Chebyshev set.*

If (c) *holds and C, x, (y_n), and y are as in the statement of* (c), *then y is the unique point of C closest to x.*

PROOF. It follows immediately from Theorem 5.3.20 that (a) $\Rightarrow$ (b), and it is obvious that (b) $\Rightarrow$ (c). Suppose next that (c) holds and that C, x, (y_n), and y are as in the statement of (c). Then $\|x - y\| = d(x, C)$. Furthermore, if y' is any member of C such that $\|x - y'\| = d(x, C)$, then (c) implies that the sequence $(y, y', y, y', \dots)$ converges, which shows that $y = y'$. This establishes the last statement in the theorem. It is now clear that (c) $\Rightarrow$ (d) since for every nonempty closed convex subset C of X and every x in X there is a minimizing sequence in C with respect to x.

Finally, suppose that (d) holds. Let (z_n) be a sequence in S_X for which there is an element z^* of S_{X^*} such that $\operatorname{Re} z^* z_n \to 1$. The goal is to show that (z_n) converges, so, after perhaps discarding some terms from the beginning of the sequence, it may be assumed that $\operatorname{Re} z^* z_n > 0$ for each n. Let $A = \{ x : x \in X, \operatorname{Re} z^* x = 1 \}$, a nonempty closed convex subset of X, and let $w_n = (\operatorname{Re} z^* z_n)^{-1} z_n$ for each n. It is clear that $d(0, A) \geq 1$, and thus that $d(0, A) = 1$ since the sequence (w_n) lies in A and $\|w_n\| \to 1$. Since every subsequence of (w_n) is a minimizing sequence in A with respect to 0, it follows that every subsequence of (w_n) has a subsequence that converges to the unique point w of A closest to 0, so $w_n \to w$. Therefore $z_n \to w$, which by Corollary 5.3.18 proves that X is a strongly rotund Banach space. This shows that (d) $\Rightarrow$ (a). ∎

The property given in part (c) of Theorem 5.3.17 might seem to have some of the flavor of the Radon-Riesz property. In fact, it implies the Radon-Riesz property.

5.3.22 Theorem. (K. Fan and I. Glicksberg, 1958 [77]). *Every strongly rotund normed space has the Radon-Riesz property. In symbols,* (K) $\Rightarrow$ (H).

PROOF. Suppose that X is a strongly rotund normed space and that (x_n) is a sequence in X and x an element of X such that $x_n \xrightarrow{w} x$ and $\|x_n\| \to \|x\|$. Since the goal is to prove that $x_n \to x$, it may be assumed that $x \neq 0$. It is easy to see that it may then be assumed that x and each x_n lie in S_X. Let x^* be a member of S_{X^*} such that $x^* x = \operatorname{Re} x^* x = 1$. Then $\operatorname{Re} x^* x_n \to 1$, so by Theorem 5.3.17 the sequence (x_n) is Cauchy. Therefore (x_n) converges in the completion of X to some y, which together with the fact that $x_n \xrightarrow{w} x$ implies that $x_n \to x$. ∎

Combining the Radon-Riesz property with a few other conditions produces a partial converse for the preceding result.

5.3.23 Theorem. (K. Fan and I. Glicksberg, 1958 [77]). *Every reflexive rotund normed space having the Radon-Riesz property is strongly rotund. In symbols,* (Rf) & (R) & (H) $\Rightarrow$ (K).

PROOF. Suppose that X is a normed space that is (Rf) & (R) & (H). Let C be a nonempty closed convex subset of X. Then C is a Chebyshev set by Corollary 5.1.19. By Theorem 5.3.21, it is enough to show that C is approximatively compact. Let (y_n) be a minimizing sequence in C with respect to some x in X and let y be the unique element of C closest to x. It is enough to show that (y_n) has a subsequence converging to y. By the reflexivity of X, there is a subsequence (y_{n_j}) of (y_n) that converges weakly to some element y' of the weakly closed set C, and $\|x - y'\| = d(x, C)$ since

$d(x,C) \leq \|x-y'\| \leq \liminf_j \|x-y_{n_j}\| = d(x,C)$. Therefore $y' = y$. Since $x - y_{n_j} \xrightarrow{w} x - y$ and $\|x-y_{n_j}\| \to \|x-y\|$, it follows that $x - y_{n_j} \to x - y$ and therefore that $y_{n_j} \to y$. ■

5.3.24 Corollary. *Every reflexive locally uniformly rotund normed space is strongly rotund. In symbols,* (Rf) & (LUR) ⇒ (K).

It follows from what has already been proved that (Rf) & (R) & (H) ⇒ (K) & (B) ⇒ (R) & (H). If it could be shown that every strongly rotund Banach space is reflexive, then a nice characterization of the strongly rotund Banach spaces would follow: They would be precisely the normed spaces that are reflexive and rotund and have the Radon-Riesz property.

This characterization does in fact hold since every strongly rotund Banach space *is* reflexive, as was first shown by Fan and Glicksberg in their 1958 paper [77]. However, this does not seem to be all that easy to prove from elementary principles. The known proofs (or at least the ones known to this author at the time of this writing) all use some form of James's theorem or, in one case, the Bishop-Phelps subreflexivity theorem; see [168] and Exercises 5.33 and 5.65. Both of these results appear in optional sections of this book, and neither is particularly trivial. Though Fan and Glicksberg published their result several years before the general case of James's theorem appeared in James's 1964 paper [112], they did use a weak form of James's theorem from his 1957 paper [111].

Though reflexive locally uniformly rotund normed spaces are strongly rotund, it follows from the remarks of the preceding paragraph that nonreflexive locally uniformly rotund Banach spaces, such as the one of Example 5.3.6, cannot be strongly rotund, so (LUR) does not imply (K). Mark Smith's paper [219] contains additional examples to show that (K) neither implies nor is implied by any of the properties (LUR), (wUR), and (wLUR).

For more on strong rotundity, see [224] and [226], and in particular Fan and Glicksberg's paper [77] in which a number of reformulations of the property are given in addition to the ones obtained above.

The last generalization of uniform rotundity to be studied in the body of this section is obtained by starting with the first sequential characterization of uniform rotundity in Proposition 5.2.8 and then stiffening the hypotheses on the sequences (x_n) and (y_n) by requiring not just that $\left\|\frac{1}{2}(x_n + y_n)\right\|$ tend to 1, but that there actually be an element of S_X to which $\frac{1}{2}(x_n + y_n)$ converges.

5.3.25 Definition. (K. W. Anderson, 1960 [5]). Suppose that X is a normed space such that whenever (x_n) and (y_n) are sequences in S_X and $\frac{1}{2}(x_n + y_n)$ converges to some member of S_X, it follows that $\|x_n - y_n\| \to 0$. Then X is *midpoint locally uniformly rotund* or *midpoint locally uniformly convex*. The abbreviation MLUR is used for this property.

Notice that if (x_n) and (y_n) are sequences in the unit sphere of a normed space X that is (MLUR), and if $\frac{1}{2}(x_n + y_n)$ converges to some z in S_X, then both x_n and y_n also converge to z. In fact, a moment's thought shows that an equivalent definition of midpoint local uniform rotundity would have been obtained if in Definition 5.3.25 it were required that x_n and y_n, or for that matter just x_n, tend to the same limit as $\frac{1}{2}(x_n + y_n)$, rather than requiring that $\|x_n - y_n\| \to 0$. There is a large number of such ways to make small modifications to the conclusion about (x_n) and (y_n) in Definition 5.3.25 without changing the property being defined. The following result gives one simple one that will be useful in what is to follow.

5.3.26 Proposition. *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is midpoint locally uniformly rotund.*

(b) *Whenever (x_n) and (y_n) are sequences in X such that $\|x_n\|$ and $\|y_n\|$ tend to 1 and $\frac{1}{2}(x_n+y_n)$ converges to some member of S_X, it follows that $\|x_n - y_n\| \to 0$.*

PROOF. All that needs to be proved is that (a) $\Rightarrow$ (b). Suppose that X is (MLUR) and that (x_n) and (y_n) are sequences in X such that $\|x_n\|$ and $\|y_n\|$ tend to 1 and $\frac{1}{2}(x_n + y_n)$ converges to some z in S_X. It may be assumed that no x_n or y_n is zero. Then

$$\begin{aligned} 0 &\le \left\|\tfrac{1}{2}\left(\|x_n\|^{-1}x_n + \|y_n\|^{-1}y_n\right) - z\right\| \\ &\le \tfrac{1}{2}\left\|\|x_n\|^{-1}x_n - x_n\right\| + \tfrac{1}{2}\left\|\|y_n\|^{-1}y_n - y_n\right\| + \left\|\tfrac{1}{2}(x_n + y_n) - z\right\| \end{aligned}$$

for each n, from which it follows that $\frac{1}{2}\left(\|x_n\|^{-1}x_n+\|y_n\|^{-1}y_n\right) \to z$. Therefore $\left\|\|x_n\|^{-1}x_n - \|y_n\|^{-1}y_n\right\| \to 0$. This and the fact that

$$\begin{aligned} 0 &\le \|x_n - y_n\| \\ &\le \left\|x_n - \|x_n\|^{-1}x_n\right\| + \left\|\|x_n\|^{-1}x_n - \|y_n\|^{-1}y_n\right\| + \left\|\|y_n\|^{-1}y_n - y_n\right\| \end{aligned}$$

for each n together show that $\|x_n - y_n\| \to 0$, as required. ■

It is clear from Proposition 5.2.8 that every uniformly rotund normed space is midpoint locally uniformly rotund. In fact, two of the other generalizations of uniform rotundity studied in this section also imply midpoint local uniform rotundity.

5.3.27 Proposition. *Every normed space that is either strongly rotund or locally uniformly rotund is midpoint locally uniformly rotund, and every midpoint locally uniformly rotund normed space is rotund. In symbols,* (K) $\Rightarrow$ (MLUR) *and* (LUR) $\Rightarrow$ (MLUR) $\Rightarrow$ (R).

PROOF. Suppose first that the normed space X is not rotund. Then there are distinct members x and y of S_X such that $\frac{1}{2}(x+y) \in S_X$, which allows the construction in the obvious way of trivial sequences (x_n) and (y_n) in S_X such that $\frac{1}{2}(x_n + y_n)$ converges to the element $\frac{1}{2}(x+y)$ of S_X even though $\|x_n - y_n\|$ does not tend to 0. Therefore X is not midpoint locally uniformly rotund.

Now suppose instead that X is strongly rotund and that (x_n) and (y_n) are sequences in S_X such that $\frac{1}{2}(x_n + y_n)$ converges to some member z of S_X. Then there is some z^* in S_{X^*} such that $z^*z = \operatorname{Re} z^*z = 1$. Since $\operatorname{Re} z^*x_n \le 1$ and $\operatorname{Re} z^*y_n \le 1$ for each n and $\frac{1}{2}(\operatorname{Re} z^*x_n + \operatorname{Re} z^*y_n) \to 1$, it follows that $\operatorname{Re} z^*x_n \to 1$ and $\operatorname{Re} z^*y_n \to 1$. Let (z_n) be the sequence $(x_1, y_1, x_2, y_2, \dots)$. Then $\operatorname{Re} z^*z_n \to 1$, so (z_n) is Cauchy by Theorem 5.3.17. Therefore $\|x_n - y_n\| \to 0$, which shows that X is (MLUR).

Finally, suppose that X is locally uniformly rotund instead of strongly rotund, and again suppose that (x_n) and (y_n) are sequences in S_X such that $\frac{1}{2}(x_n + y_n)$ converges to some z in S_X. Since

$$2 \ge \left\|\tfrac{1}{2}(x_n + z)\right\| + 1 \ge \left\|\tfrac{1}{2}(x_n + y_n) + z\right\| \to 2$$

as $n \to \infty$, it follows that $\left\|\frac{1}{2}(x_n + z)\right\| \to 1$. Therefore $\|x_n - z\| \to 0$ by Proposition 5.3.5. Similarly, it follows that $\|y_n - z\| \to 0$, so $\|x_n - y_n\| \to 0$. This establishes that X is (MLUR). ■

As was mentioned above, it is known that every strongly rotund Banach space is reflexive. Since the Banach space of Example 5.3.6 is (MLUR) but not reflexive, it follows that (MLUR) does not imply (K). Smith's paper [219] contains other examples to show that (R) does not imply (MLUR) and that (MLUR) does not imply (LUR). His paper also has examples to show that (MLUR) neither implies nor is implied by either (wUR) or (wLUR). Smith later settled the question of whether midpoint local uniform rotundity implies the Radon-Riesz property by equivalently renorming c_0 to be (MLUR) but not (H); see his 1981 paper [220].

Midpoint locally uniformly rotund normed spaces have a characterization in terms of approximative compactness analogous to the one for strongly rotund Banach spaces given in Theorem 5.3.21.

5.3.28 Theorem. (R. E. Megginson, 1984 [168]). *A normed space X is midpoint locally uniformly rotund if and only if every closed ball in X is an approximatively compact Chebyshev set.*

PROOF. Suppose first that X is (MLUR), that x is a point in X, that C is a closed ball in X with center c and radius r, and that (y_n) is a minimizing sequence in C with respect to x. The immediate goal is to prove that there is a unique point of C closest to x and that (y_n) has a subsequence convergent to a member of C, so it may be assumed that $x = 0$ and that $d(0, C) = 1$. First of all, there cannot be two different points of C at distance 1 from 0, for it would then follow from the convexity of $C \cap B_X$ that the entire straight line segment connecting the two points would lie in S_X, a contradiction to the rotundity of X. Let y be the point where the straight line segment connecting 0 to c intersects the boundary of C. Then the triangle inequality implies that y is the unique point of C closest to 0. Notice that $\|y\| = d(0, C) = 1$ and that $c = (1 + r)y$.

It will now be shown that (y_n) has a subsequence converging to an element of C, and in fact that $y_n \to y$.

Case 1: $r = 1$. For each positive integer n, let $y'_n = 2y - y_n$, and observe that $\frac{1}{2}(y_n + y'_n) = y$. Since $c = 2y$,

$$2 = \|2y\| \leq \|2y - y_n\| + \|y_n\| \leq 1 + \|y_n\|$$

for each n, so $\|y'_n\| = \|2y - y_n\| \to 1$ because $\|y_n\| \to 1$. It then follows from Proposition 5.3.26 that $\|y_n - y'_n\| \to 0$, so $y_n \to y$.

Case 2: $r < 1$. In this case $\|c - 2y\| = \|(r-1)y\| = 1 - r$. For each member z of C,

$$\|z - 2y\| \leq \|z - c\| + \|c - 2y\| \leq r + 1 - r = 1,$$

which shows that C is included in the closed ball of radius 1 centered at $2y$. Therefore (y_n) is a minimizing sequence in this larger ball with respect to 0, so all that is needed to show that $y_n \to y$ is an appeal to case 1.

Case 3: $r > 1$. Let C' be the closed ball of radius 1 centered at $2y$. For each positive integer n, let $z_n = y + r^{-1}(y_n - y)$ and observe that $z_n \in C'$ because

$$\|z_n - 2y\| = \|r^{-1}(y_n - y) - y\| = r^{-1}\|y_n - (1 + r)y\| = r^{-1}\|y_n - c\| \leq 1.$$

Since

$$1 \leq \|z_n\| \leq \|(1 - r^{-1})y\| + \|r^{-1}y_n\| = (1 - r^{-1}) + r^{-1}\|y_n\|$$

for each n, and since $(1 - r^{-1}) + r^{-1}\|y_n\| \to 1$, it follows that $\|z_n\| \to 1$ and therefore, by case 1, that $z_n \to y$. Therefore $y_n \to y$.

In any case, there is a unique point of C closest to 0, and the minimizing sequence (y_n) in C with respect to 0 converges to the element y of C, so every closed ball in X is an approximatively compact Chebyshev set.

Now suppose conversely that every closed ball in X is an approximatively compact Chebyshev set. Let (u_n) and (v_n) be sequences in S_X such that there is an element w of S_X to which $\frac{1}{2}(u_n + v_n)$ converges. Then $d(2w, B_X) = 1$ and w is the unique point of B_X closest to $2w$. Since

$$1 \leq \|u_n - 2w\| \leq \|u_n + v_n - 2w\| + \|v_n\| = \|u_n + v_n - 2w\| + 1$$

for each n and since $\|u_n + v_n - 2w\| \to 0$, each subsequence of (u_n) is a minimizing sequence in B_X with respect to $2w$ and therefore must have a convergent subsequence whose limit is necessarily the unique point w of B_X closest to $2w$. It follows that $u_n \to w$, and similarly that $v_n \to w$, and therefore that $\|u_n - v_n\| \to 0$. This proves that X is (MLUR). ■

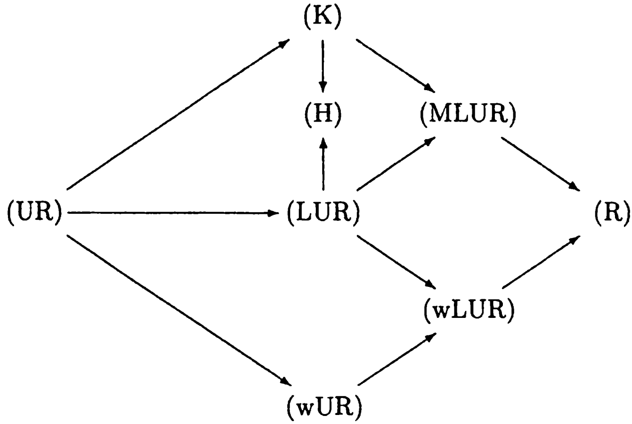


FIGURE 5.3. An implication diagram for generalizations of uniform rotundity.

Figure 5.3 shows all of the implications between single properties proved in this section, except for those involving weak* uniform rotundity that apply only to dual spaces; however, see also Theorem 5.3.23 and Corollary 5.3.24. Notice that (H) does not imply any of the other properties in the diagram, since ℓ_1^2 is not rotund but, like all finite-dimensional normed spaces, has the Radon-Riesz property. It follows from this and Smith's examples already mentioned above that no arrows can be added to this diagram showing relationships not already implied by the diagram.

Extensive lists of other generalizations of uniform rotundity as well as some of their characteristics and the relationships between them can be found in [56, pp. 145–147] and [107, pp. 71–93]. A few of these generalizations can be found in the exercises for this section. Two of the generalizations from the exercises, namely, uniform rotundity in weakly compact sets of directions and uniform rotundity in every direction, as well as weak uniform rotundity and uniform rotundity itself, are *directionalizations* of uniform rotundity in the sense that they can be defined in terms of the following modulus. See Exercises 5.29–5.32.

5.3.29 Definition. Suppose that X is a normed space. Define a function $\delta_X\colon [0,2] \times (X \setminus \{0\}) \to [0,1]$ by the formula

$$\delta_X(\epsilon, \to z) = \inf\Big\{\, 1 - \Big\|\tfrac{1}{2}(x+y)\Big\| : x, y \in S_X,\ \|x-y\| \geq \epsilon,\ x - y = \alpha z \text{ for some scalar } \alpha \,\Big\}.$$

Then δ_X is the *directional modulus of rotundity* of X. If A is a nonempty subset of $X \setminus \{0\}$, then

$$\delta_X(\epsilon, \to A) = \inf\{\, \delta_X(\epsilon, \to z) : z \in A \,\}$$

whenever $0 \leq \epsilon \leq 2$.

A good starting point for discovering more about approximative compactness and other approximative properties of sets in normed spaces is L. P. Vlasov's 1973 survey article [238]. Those interested in the subject of equivalently renorming spaces to have desired properties should read Mark Smith's articles [219] and [220] to learn about the art from one of its masters.

Exercises

5.24 Prove the parallelogram law for ℓ_2 mentioned in Example 5.3.6.

5.25 Give at least one sequential characterization of weak local uniform rotundity.

5.26 Find a sequential characterization of weak* uniform rotundity.

5.27 (a) Prove that every approximatively compact set is closed.

(b) Give an example of an approximatively compact set that is not compact, or even weakly compact.

5.28 (K. Fan and I. Glicksberg, 1958 [77]). Suppose that k is a positive integer such that $k \geq 2$. A normed space X is *k-rotund* if each sequence (x_n) in X such that $\lim_{n_1,\ldots,n_k\to\infty}\|k^{-1}\sum_{j=1}^{k} x_{n_j}\| = 1$ is Cauchy. (The notation $\lim_{n_1,\ldots,n_k\to\infty}\|k^{-1}\sum_{j=1}^{k} x_{n_j}\| = 1$ means that for every positive ϵ there is a positive integer N_ϵ such that $\big|\|k^{-1}\sum_{j=1}^{k} x_{n_j}\| - 1\big| < \epsilon$ whenever $n_1, \ldots, n_k \geq N_\epsilon$.) The abbreviation kR is used for the property of k-rotundity.

(a) Prove that every uniformly rotund normed space is k-rotund, and that every k-rotund normed space is $(k+1)$-rotund. That is, show that (UR) $\Rightarrow$ (kR) $\Rightarrow$ $\big((k+1)\text{R}\big)$.

(b) Prove that every k-rotund normed space is strongly rotund. That is, show that (kR) $\Rightarrow$ (K).

5.29 Prove that a normed space X is (UR) if and only if $\delta_X(\epsilon, \to A) > 0$ whenever $0 < \epsilon \leq 2$ and A is a nonempty bounded closed subset of $X \setminus \{0\}$.

5.30 Prove that a normed space X is (wUR) if and only if $\delta_X(\epsilon, \to A) > 0$ whenever $0 < \epsilon \leq 2$ and A is a nonempty bounded weakly closed subset of $X \setminus \{0\}$.

5.31 (M. A. Smith, 1975, 1977 [217, 218]). A normed space X is *uniformly rotund in weakly compact sets of directions* if $\delta_X(\epsilon, \to A) > 0$ whenever $0 < \epsilon \leq 2$ and A is a nonempty weakly compact subset of $X \setminus \{0\}$. The abbreviation URWC is used for this property.

(a) Prove that a normed space X is (URWC) if and only if it has this property: Whenever (x_n) and (y_n) are sequences in S_X such that $\|\frac{1}{2}(x_n + y_n)\| \to 1$ and $x_n - y_n \xrightarrow{w} v$ for some v in X, it follows that $v = 0$.

(b) Show that (wUR) $\Rightarrow$ (URWC) $\Rightarrow$ (R).

(c) Show that (Rf) & (wUR) $\Leftrightarrow$ (Rf) & (URWC); that is, that weak uniform rotundity and uniform rotundity in weakly compact sets of directions are equivalent properties for reflexive normed spaces.

Smith has shown in [219] that the nonreflexive Banach space $(\ell_1, \|\cdot\|_E)$ of Example 5.3.6 is (URWC) but not (wUR).

5.32 (A. L. Garkavi, 1962 [84]). A normed space X is *uniformly rotund in every direction* or *uniformly convex in every direction* or *directionally uniformly rotund* if $\delta_X(\epsilon, \rightarrow z) > 0$ whenever $0 < \epsilon \leq 2$ and $z \in S_X$. The abbreviation URED is used for this property.

(a) Prove that a normed space X is (URED) if and only if it has this property: Whenever (x_n) and (y_n) are sequences in S_X such that $\|\frac{1}{2}(x_n + y_n)\| \rightarrow 1$ and such that $x_n - y_n \in \langle\{v\}\rangle$ for some v in X and each n, it follows that $x_n - y_n \rightarrow 0$.

(b) Show that (wUR) $\Rightarrow$ (URED) $\Rightarrow$ (R).

(c) (This uses material from Exercise 5.31). Show that (URWC) $\Rightarrow$ (URED).

Smith gave an example in [219] of a Banach space that is (URED) but not (URWC), and another Banach space that is (R) but not (URED); both examples are formed by equivalently renorming ℓ_2. It can be shown that a normed space X is (URED) if and only if $\delta_X(\epsilon, \rightarrow A) > 0$ whenever $0 < \epsilon \leq 2$ and A is a nonempty compact subset of $X \setminus \{0\}$; see [218]. Another good source of information on uniform rotundity in every direction is the paper of Day, James, and Swaminathan [57] devoted to the property.

5.33 This exercise requires James's theorem from either of the optional Sections 1.13 and 2.9. Prove that a normed space is a strongly rotund Banach space if and only if it is a reflexive rotund normed space with the Radon-Riesz property. That is, show that (K) & (B) $\Leftrightarrow$ (R) & (Rf) & (H).

5.34 (Ivan Singer, 1964 [215]). A normed space X has the *Efimov-Stechkin property* if, whenever (x_n) is a sequence in S_X for which there is an element x^* of S_{X^*} such that $\operatorname{Re} x^* x_n \rightarrow 1$, the sequence (x_n) has a convergent subsequence. The abbreviation CD is used for this property.

(a) Show that a normed space X has the Efimov-Stechkin property if and only if every nonempty closed convex subset of X is approximatively compact.

(b) Prove that every strongly rotund Banach space has the Efimov-Stechkin property.

(c) Show by example that not every Banach space with the Efimov-Stechkin property is rotund.

(d) Prove that every normed space with the Efimov-Stechkin property is a Banach space.

(e) Prove that every normed space with the Efimov-Stechkin property has the Radon-Riesz property.

(f) The rest of this exercise requires James's theorem from either of the optional Sections 1.13 and 2.9. Prove that every normed space with the Efimov-Stechkin property is reflexive.

(g) Prove that a normed space has the Efimov-Stechkin property if and only if it is reflexive and has the Radon-Riesz property. That is, show that (CD) $\Leftrightarrow$ (Rf) & (H).

(h) Conclude from the results of this exercise and Exercise 5.33 that a normed space is a strongly rotund Banach space if and only if it is rotund and has the Efimov-Stechkin property. That is, show that (K) & (B) $\Leftrightarrow$ (R) & (CD).

5.35 The purpose of this exercise is to present a technique that can be used to design a real normed space having $\mathbb{R}^2$ as its underlying vector space such that the closed unit ball of the space has certain specified geometric properties.

(a) Prove that if a subset C of $\mathbb{R}^2$ is balanced (which in this case is equivalent to being symmetric about the origin), convex, and closed, and has nonempty interior with respect to the Euclidean topology for $\mathbb{R}^2$, then C is the closed unit ball for some norm on $\mathbb{R}^2$ that is equivalent to the Euclidean norm of $\mathbb{R}^2$.

(b) Use (a) to construct a Banach space X with the property that for some element x of S_X there can be found a different element y of S_X such that $\frac{1}{2}(x+y) \in S_X$, but for some other element x' of S_X there is no y' in S_X such that $x' \neq y'$ and $\frac{1}{2}(x'+y') \in S_X$.

5.36 (R. C. James, 1964 [113]). Suppose that X is a normed space and that δ_X is the modulus of rotundity of X. Then X is *inquadrate* or *uniformly nonsquare* if there exists an ϵ such that $0 < \epsilon < 2$ and $\delta_X(\epsilon) > 0$. The abbreviation NQ is used for this property. It is obvious that this property is implied by uniform rotundity. Show that it does *not* imply rotundity. The method of Exercise 5.35 might be helpful.

James has shown in [113] that every inquadrate Banach space is reflexive. Notice that it follows from this and Exercise 5.14 that if δ_Y is the modulus of rotundity for a rotund nonreflexive Banach space Y, then

$$\delta_Y(\epsilon) = \begin{cases} 0 & \text{if } 0 < \epsilon < 2; \\ 1 & \text{if } \epsilon = 2. \end{cases}$$

It even turns out that a normed space is $\langle \mathrm{NQ} \rangle$ if and only if it is $\langle \mathrm{UR} \rangle$, that is, if and only if it is superreflexive. See [56, pp. 169–173] for a proof of this and for references concerning the history of the result.

5.4 Smoothness

In the introduction to this chapter, it was said that a normed space's closed unit ball is *smooth* if the unit sphere has no "corners" or "sharp bends." It is time to make this "definition" a bit more rigorous. One clue as to how this could be done is given by the illustration of the unit sphere of real ℓ_1^2 in Figure 5.4. Through each of the four corners of this unit sphere, it is

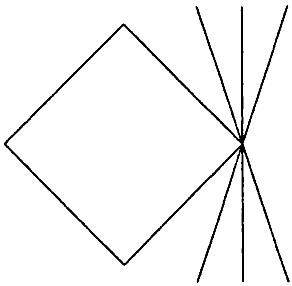


FIGURE 5.4. Multiple support hyperplanes for $B_{\ell_1^2}$ at a point of $S_{\ell_1^2}$.

possible to pass more than one line that does not penetrate the interior of the closed unit ball, while this is not possible at any of the other points of $S_{\ell_1^2}$. Since the support hyperplanes for $B_{\ell_1^2}$ are precisely the straight lines in ℓ_1^2 that intersect $S_{\ell_1^2}$ but not $B_{\ell_1^2}^\circ$, this suggests the following definition.

5.4.1 Definition. Suppose that x_0 is an element of the unit sphere of a normed space X. Then x_0 is a *point of smoothness* of B_X if there is no more than one support hyperplane for B_X that supports B_X at x_0. The space X is *smooth* if each point of S_X is a point of smoothness of B_X.

By Corollary 1.9.8, each point x_0 of the unit sphere of a normed space X is a support point for B_X and therefore gives rise to at least one support hyperplane for B_X that supports B_X at x_0. It would therefore be equivalent to replace "no more than one support hyperplane" in the preceding definition by "exactly one support hyperplane."

If the closed unit ball of a normed space X is supported at some point x_0 of S_X by elements x_1^* and x_2^* of S_{X^*}, and if x_1^* and x_2^* induce the same support hyperplane H for B_X, then

$$H = \{\, x : x \in X,\ \operatorname{Re} x_1^* x = 1 \,\} = \{\, x : x \in X,\ \operatorname{Re} x_2^* x = 1 \,\},$$

from which it follows that $\operatorname{Re} x_1^* = \operatorname{Re} x_2^*$ and therefore that $x_1^* = x_2^*$. The following characterizations of points of smoothness and the smoothness property follow easily.

5.4.2 Proposition. *Suppose that X is a normed space and that $x_0 \in S_X$. Then the following are equivalent.*

(a) *The element x_0 is a point of smoothness of B_X.*

(b) *There is exactly one element x_0^* of S_{X^*} that supports B_X at x_0.*

(c) *There is exactly one element x_0^* of S_{X^*} such that $\operatorname{Re} x_0^* x_0 = 1$.*

(d) *There is exactly one element x_0^* of S_{X^*} such that $x_0^* x_0 = 1$.*

5.4.3 Corollary. *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is smooth.*

(b) *For each x in S_X, there is a unique x^* in S_{X^*} that supports B_X at x.*

(c) *For each x in S_X, there is a unique x^* in S_{X^*} such that $\operatorname{Re} x^*x = 1$.*

(d) *For each x in S_X, there is a unique x^* in S_{X^*} such that $x^*x = 1$.*

The following result is essentially obvious, and in any case is an easy consequence of the preceding corollary.

5.4.4 Proposition. *Every normed space that is isometrically isomorphic to a smooth normed space is itself smooth.*

This does not generalize to isomorphisms. For example, real ℓ_2^2 is obviously smooth even though it is isomorphic to the nonsmooth normed space ℓ_1^2.

As will be apparent from the examples given below, the classical Banach spaces studied earlier in this book tend to be both rotund and smooth when they have either property. However, neither property actually implies the other; see Exercise 5.37. It is true that the presence of either one in a normed space is implied by the existence of the other in the dual space.

5.4.5 Proposition. *A normed space is smooth if its dual space is rotund.*

PROOF. Suppose that X is a nonsmooth normed space. Then for some x in S_X there are two different elements x_1^* and x_2^* of S_{X^*} such that $x_1^*x = x_2^*x = 1$. Therefore $\frac{1}{2}(x_1^* + x_2^*)(x) = 1$, so $1 \le \left\|\frac{1}{2}(x_1^* + x_2^*)\right\| \le 1$, and thus $\left\|\frac{1}{2}(x_1^* + x_2^*)\right\| = 1$. It follows that X^* is not rotund. ■

5.4.6 Proposition. *A normed space is rotund if its dual space is smooth.*

PROOF. Suppose that X is a nonrotund normed space and that Q is the natural map from X into X^{**}. Then there are two different elements x_1 and x_2 of S_X and an element x^* of S_{X^*} such that $1 = x^*\left(\frac{1}{2}(x_1 + x_2)\right) = x^*x_1 = x^*x_2 = (Qx_1)(x^*) = (Qx_2)(x^*)$. The space X^* is therefore not smooth. ■

An obvious argument based on the preceding three propositions then yields the following result.

5.4.7 Proposition. *A reflexive normed space is rotund if and only if its dual space is smooth, and is smooth if and only if its dual space is rotund.*

The implications in Propositions 5.4.5 and 5.4.6 cannot in general be reversed for nonreflexive normed spaces. The rotund Banach space $\ell_{1,r}$ of

Example 5.1.8 obtained by equivalently renorming ℓ_1 has its dual space isomorphic to ℓ_∞, but it was shown by M. M. Day in a 1955 paper [51] that ℓ_∞ cannot be equivalently renormed to be smooth. See Exercise 5.41 for another proof that $\ell_{1,r}^*$ is not smooth. Also, see S. L. Troyanski's 1970 paper [231] for an example of a smooth Banach space whose dual is not rotund, and [107, pp. 130–132] for an English exposition of the example.

5.4.8 Example. The scalar field $\mathbb{F}$, viewed as a normed space over $\mathbb{F}$, is obviously smooth. The same holds for every zero- or one-dimensional normed space.

5.4.9 Example. Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω and that $1 < p < \infty$. Let q be the real number such that $p^{-1} + q^{-1} = 1$. Then $\big(L_p(\Omega, \Sigma, \mu)\big)^*$ is isometrically isomorphic to the rotund normed space $L_q(\Omega, \Sigma, \mu)$, so $L_p(\Omega, \Sigma, \mu)$ is smooth.

5.4.10 Example. As in the last example, suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω, but suppose also that there are disjoint measurable subsets A_1 and A_2 of Ω each with finite positive measure. Let $\mathbf{I}_{A_1}$ and $\mathbf{I}_{A_2}$ be the respective indicator functions of these sets, and let $f_1 = \big(\mu(A_1)\big)^{-1}\mathbf{I}_{A_1}$, $f_2 = \big(\mu(A_2)\big)^{-1}\mathbf{I}_{A_2}$, $g_1 = \mathbf{I}_{A_1} + \mathbf{I}_{A_2}$, and $g_2 = \mathbf{I}_{A_1} - \mathbf{I}_{A_2}$. Then g_1 and g_2 can be identified in the usual way with different norm-one bounded linear functionals on $L_1(\Omega, \Sigma, \mu)$, and similarly f_1 and f_2 can be identified with different norm-one bounded linear functionals on $L_\infty(\Omega, \Sigma, \mu)$. Notice that this can be proved directly; it does not require that $\big(L_1(\Omega, \Sigma, \mu)\big)^*$ be identifiable with $L_\infty(\Omega, \Sigma, \mu)$ through some assumption such as that the measure space is σ-finite. Since

$$\int_\Omega f_1 g_2 = \int_\Omega f_1 g_1 = \int_\Omega f_2 g_1 = 1,$$

neither $L_1(\Omega, \Sigma, \mu)$ nor $L_\infty(\Omega, \Sigma, \mu)$ is smooth. In particular, the spaces ℓ_1 and ℓ_∞ are not smooth, nor are ℓ_1^n and ℓ_∞^n when $n \geq 2$. Notice that the spaces ℓ_1^n and ℓ_∞^n are smooth when $n < 2$ by Example 5.4.8.

5.4.11 Example. Let x_1^* and x_2^* be the members of $S_{c_0^*}$ identified in the usual way with the first two standard unit vectors of ℓ_1, and let x be the member $(1, 1, 0, 0, 0, \dots)$ of S_{c_0}. Then $x_1^* x = x_2^* x = 1$, so c_0 is not smooth. Since x_1^* and x_2^* can be viewed in the obvious way as members of $S_{\ell_\infty^*}$, it also follows that ℓ_∞ is not smooth.

5.4.12 Example. Suppose that K is a compact Hausdorff space with more than one element. Let k_1 and k_2 be different members of K. For $j = 1, 2$, let δ_j be the member of $S_{\mathrm{rca}(K)}$ defined by the formula

$$\delta_j(A) = \begin{cases} 1 & \text{if } k_j \in A; \\ 0 & \text{if } k_j \notin A. \end{cases}$$

Let f be the member of $S_{C(K)}$ that takes on the value 1 everywhere on K. Since δ_1 and δ_2 can be viewed in the usual way as norm-one bounded linear functionals on $C(K)$, and since

$$\int_K f\,d\delta_1 = \int_K f\,d\delta_2 = 1,$$

the normed space $C(K)$ is not smooth.

5.4.13 Example. This is an example of a nonreflexive smooth Banach space. Define $T\colon \ell_2 \to c_0$ by the formula $T\big((\alpha_n)\big) = (\alpha_n)$. Then T is a one-to-one bounded linear map having norm 1, and $T(\ell_2)$ is dense in c_0 since $T(\ell_2)$ includes the vector space of finitely nonzero sequences. Therefore the adjoint T^* of T is a one-to-one weakly* continuous bounded linear map from c_0^* into the rotund normed space ℓ_2^*, and $\|T^*\| = 1$. Let $\|x^*\|_a = \|x^*\| + \|T^*x^*\|$ for each x^* in c_0^*. Then $\|x^*\| \le \|x^*\|_a \le 2\|x^*\|$ whenever $x^* \in c_0^*$, from which it follows easily that $\|\cdot\|_a$ is a norm on c_0^* equivalent to its original dual norm.

Suppose that $x_1^*, x_2^* \in c_0^*$ and that $\|x_1^* + x_2^*\|_a = \|x_1^*\|_a + \|x_2^*\|_a$. Then $\|Tx_1^* + Tx_2^*\| = \|Tx_1^*\| + \|Tx_2^*\|$, so it follows from the rotundity of ℓ_2^* that one of the two vectors Tx_1^* and Tx_2^* is a nonnegative multiple of the other. Since T^* is one-to-one, this implies that one of the vectors x_1^* and x_2^* must be a nonnegative multiple of the other, so $\|\cdot\|_a$ is a rotund norm.

It follows from the weak* lower semicontinuity of the original norms of c_0^* and ℓ_2^* and the weak*-to-weak* continuity of T^* that $\|w^*\text{-}\lim_\alpha x_\alpha^*\|_a \le \liminf_\alpha \|x_\alpha^*\|_a$ whenever (x_α^*) is a weakly* convergent net in c_0^*. By Theorem 2.6.15, there is a norm $\|\cdot\|_b$ on c_0 equivalent to its original norm such that $\|\cdot\|_a$ is the dual norm on $(c_0, \|\cdot\|_b)^*$. Since $\|\cdot\|_a$ is a rotund norm, the Banach space $(c_0, \|\cdot\|_b)$ is smooth, but is not reflexive since it is isomorphic to c_0.

Just as there is a connection between the smoothness of the graph of a real-valued function of a real variable and the differentiability of the function, so is there a connection between the smoothness of the unit sphere of a normed space and the *Gateaux differentiability* of the norm. In the interest of sticking to the subject at hand, the following treatment of Gateaux differentiability is done specifically for norm functions. A far more general and extensive discussion of Gateaux differentiability can be found in [88].

5.4.14 Lemma. *Suppose that X is a normed space and that $x_0 \in S_X$. Then for each y in X, the function*[3]

$$t \mapsto \frac{\|x_0 + ty\| - \|x_0\|}{t}$$

---

[3]The reason for not substituting 1 for $\|x_0\|$ in the following quotient or in similar quotients in the rest of this chapter is to make the point that these expressions represent difference quotients for the norm function.

from $\mathbb{R} \setminus \{0\}$ into $\mathbb{R}$ is nondecreasing, so both

$$\lim_{t \to 0^-} \frac{\|x_0 + ty\| - \|x_0\|}{t}$$

and

$$\lim_{t \to 0^+} \frac{\|x_0 + ty\| - \|x_0\|}{t}$$

exist, and

$$\lim_{t \to 0^-} \frac{\|x_0 + ty\| - \|x_0\|}{t} \leq \lim_{t \to 0^+} \frac{\|x_0 + ty\| - \|x_0\|}{t}.$$

Furthermore, the function

$$y \mapsto \lim_{t \to 0^+} \frac{\|x_0 + ty\| - \|x_0\|}{t}$$

is a sublinear functional on X.

PROOF. For the moment, consider the element y of X to be fixed, and let

$$f(t) = \frac{\|x_0 + ty\| - \|x_0\|}{t}$$

whenever $t \in \mathbb{R} \setminus \{0\}$. Suppose that $0 < t_1 < t_2$. Then

$$\begin{aligned}
\|x_0 + t_1 y\| - \|x_0\| &= \left\| \frac{t_1}{t_2}(x_0 + t_2 y) + \left(1 - \frac{t_1}{t_2}\right) x_0 \right\| - \|x_0\| \\
&\leq \frac{t_1}{t_2}\|x_0 + t_2 y\| + \left(1 - \frac{t_1}{t_2}\right)\|x_0\| - \|x_0\| \\
&= \frac{t_1}{t_2}\big(\|x_0 + t_2 y\| - \|x_0\|\big),
\end{aligned}$$

so

$$\frac{\|x_0 + t_1 y\| - \|x_0\|}{t_1} \leq \frac{\|x_0 + t_2 y\| - \|x_0\|}{t_2}$$

and

$$\begin{aligned}
\frac{\|x_0 - t_2 y\| - \|x_0\|}{-t_2} &= -\frac{\|x_0 + t_2(-y)\| - \|x_0\|}{t_2} \\
&\leq -\frac{\|x_0 + t_1(-y)\| - \|x_0\|}{t_1} \\
&= \frac{\|x_0 - t_1 y\| - \|x_0\|}{-t_1}.
\end{aligned}$$

Since $\|x_0\| \leq \frac{1}{2}\big(\|x_0 - t_1 y\| + \|x_0 + t_1 y\|\big)$, it is also true that

$$\frac{\|x_0 - t_1 y\| - \|x_0\|}{-t_1} \leq \frac{\|x_0 + t_1 y\| - \|x_0\|}{t_1}.$$

It follows from all of this that f is nondecreasing, so $\lim_{t\to 0^-} f(t)$ and $\lim_{t\to 0^+} f(t)$ both exist and $\lim_{t\to 0^-} f(t) \le \lim_{t\to 0^+} f(t)$.

Now drop the assumption that y is fixed, and let

$$g(y) = \lim_{t\to 0^+} \frac{\|x_0 + ty\| - \|x_0\|}{t}$$

whenever $y \in X$. Suppose that $s > 0$ and that $y_1, y_2 \in X$. Then

$$g(sy_1) = \lim_{t\to 0^+} \frac{\|x_0 + tsy_1\| - \|x_0\|}{t} = s \lim_{t\to 0^+} \frac{\|x_0 + tsy_1\| - \|x_0\|}{ts} = sg(y_1).$$

The function g is therefore positive-homogeneous. Also, if $t > 0$ then

$$\frac{\|x_0 + t(y_1 + y_2)\| - \|x_0\|}{t} \le \frac{\|x_0 + 2ty_1\| - \|x_0\|}{2t} + \frac{\|x_0 + 2ty_2\| - \|x_0\|}{2t},$$

so letting t tend to 0 through positive values shows that $g(y_1 + y_2) \le g(y_1) + g(y_2)$. Therefore g is finitely subadditive and so is sublinear. ■

5.4.15 Definition. Suppose that X is a normed space, that $x_0 \in S_X$, and that $y_0 \in X$. Let

$$G_-(x_0, y_0) = \lim_{t\to 0^-} \frac{\|x_0 + ty_0\| - \|x_0\|}{t}$$

and

$$G_+(x_0, y_0) = \lim_{t\to 0^+} \frac{\|x_0 + ty_0\| - \|x_0\|}{t}.$$

Then $G_-(x_0, y_0)$ and $G_+(x_0, y_0)$ are, respectively, the *left-hand* and *right-hand Gateaux derivative of the norm at x_0 in the direction y_0*. The norm is *Gateaux differentiable at x_0 in the direction y_0* if $G_-(x_0, y_0) = G_+(x_0, y_0)$, in which case the common value of $G_-(x_0, y_0)$ and $G_+(x_0, y_0)$ is denoted by $G(x_0, y_0)$ and is called the *Gateaux derivative of the norm at x_0 in the direction y_0*. If the norm is Gateaux differentiable at x_0 in every direction y, then the norm is *Gateaux differentiable at x_0*. Finally, if the norm is Gateaux differentiable at every point of the unit sphere, then it is simply said that the norm is *Gateaux differentiable.*

Notice that the norm of a normed space X is Gateaux differentiable at a point x_0 of the unit sphere in a direction y_0 if and only if

$$\lim_{t\to 0} \frac{\|x_0 + ty_0\| - \|x_0\|}{t}$$

exists, in which case $G(x_0, y_0)$ is the limit. Therefore the norm of the space is Gateaux differentiable if and only if the above limit exists for each x_0

in the unit sphere and each y_0 in the space. Notice also that if $x_0 \in S_X$, then $G_-(x_0, y) = -G_+(x_0, -y)$ whenever $y \in X$, which together with the sublinearity of $G_+(x_0, \cdot)$ implies that $G_-(x_0, \cdot)$ is positive-homogeneous.

5.4.16 Lemma. *Suppose that X is a normed space, that $x_0 \in S_X$, and that $x_0^* \in S_{X^*}$. Then x_0^* supports B_X at x_0 if and only if*

$$G_-(x_0, y) \le \operatorname{Re} x_0^* y \le G_+(x_0, y)$$

whenever $y \in X$.

PROOF. Suppose first that x_0^* supports B_X at x_0 and that $y \in X$. If $t > 0$, then

$$\begin{aligned} \operatorname{Re} x_0^*(ty) &= \operatorname{Re} x_0^*(x_0 + ty) - \operatorname{Re} x_0^* x_0 \\ &= \operatorname{Re} x_0^*(x_0 + ty) - \|x_0\| \\ &\le \|x_0 + ty\| - \|x_0\|, \end{aligned}$$

and so

$$\begin{aligned} \frac{\|x_0 - ty\| - \|x_0\|}{-t} &= -\frac{\|x_0 + t(-y)\| - \|x_0\|}{t} \\ &\le -\operatorname{Re} x_0^*(-y) \\ &= \operatorname{Re} x_0^* y \\ &\le \frac{\|x_0 + ty\| - \|x_0\|}{t}, \end{aligned}$$

from which it follows that $G_-(x_0, y) \le \operatorname{Re} x_0^* y \le G_+(x_0, y)$.

For the converse, suppose that $G_-(x_0, y) \le \operatorname{Re} x_0^* y \le G_+(x_0, y)$ when $y \in Y$. Since the norm is Gateaux differentiable at x_0 in the direction x_0 and $G(x_0, x_0) = 1$, it follows that $\operatorname{Re} x_0^* x_0 = 1$ and therefore that x_0^* supports B_X at x_0. ■

5.4.17 Theorem. (S. Banach, 1932 [13]). *Suppose that X is a normed space and that $x_0 \in S_X$. Then x_0 is a point of smoothness of B_X if and only if the norm of X is Gateaux differentiable at x_0. Furthermore, if x_0 is a point of smoothness of B_X and x_0^* is the unique member of S_{X^*} that supports B_X at x_0, then the Gateaux derivative of the norm of X at x_0 in each direction y is given by the formula $G(x_0, y) = \operatorname{Re} x_0^* y$.*

PROOF. Suppose first that the norm of X is not Gateaux differentiable at x_0. Then there is a y_0 in X such that $G_-(x_0, y_0) < G_+(x_0, y_0)$. Fix a real number s such that $G_-(x_0, y_0) \le s \le G_+(x_0, y_0)$, and let $f_s(ry_0) = rs$ for each real number r. Then f_s is a real-linear functional on the real vector space V consisting of all real multiples of y_0, and $f_s(ry_0) \le G_+(x_0, ry_0)$ whenever $r \in \mathbb{R}$; this uses the positive-homogeneity of $G_+(x_0, \cdot)$ and the fact that $G_+(x_0, ry_0) = -G_-(x_0, -ry_0)$ for each real r and in particular for each negative r. By the vector space version of the Hahn-Banach extension theorem, there is a linear functional x_s^* on X such that the restriction of $\operatorname{Re} x_s^*$ to V is f_s and $\operatorname{Re} x_s^* y \le G_+(x_0, y)$ whenever $y \in X$. Notice that this implies that

$$\operatorname{Re} x_s^* y = -\operatorname{Re} x_s^*(-y) \ge -G_+(x_0, -y) = G_-(x_0, y)$$

for each y in X. It also follows that

$$\operatorname{Re} x_s^* y \le G_+(x_0, y) \le \frac{\|x_0 + 1y\| - \|x_0\|}{1} \le \|y\|$$

whenever $y \in X$, so x_s^* is bounded and has norm at most 1. The norm of x_s^* is in fact 1 since $\operatorname{Re} x_s^* x_0 = G(x_0, x_0) = 1$. Therefore x_s^* is a member of S_{X^*} that supports B_X at x_0. Now s can be any one of the infinitely many real numbers in the interval $[G_-(x_0, y_0), G_+(x_0, y_0)]$, and each such s gives rise to a different x_s^* in S_{X^*} supporting B_X at x_0 since $\operatorname{Re} x_s^* y_0 = s$. Therefore x_0 is not a point of smoothness of B_X.

Now suppose conversely that the norm of X is Gateaux differentiable at x_0 and that x_0^* is a member of S_{X^*} that supports B_X at x_0. It follows from the preceding lemma that $\operatorname{Re} x_0^* y = G(x_0, y)$ whenever $y \in X$, so every other member of S_{X^*} that supports B_X at x_0 has the same real part as x_0^* and thus equals x_0^*. Therefore x_0 is a point of smoothness of B_X by Proposition 5.4.2, which finishes the proof of the equivalence that is the first conclusion of this theorem. The argument given at the beginning of this paragraph then produces the formula for the Gateaux derivative that is the other conclusion. ■

5.4.18 Corollary. *A normed space is smooth if and only if its norm is Gateaux differentiable.*

There is another characterization of points of smoothness that is historically related to Gateaux differentiability. Šmulian first stated the following result as a characterization of points at which the norm is Gateaux differentiable rather than as a characterization of points of smoothness.

5.4.19 Theorem. (V. L. Šmulian, 1939 [222]). *Suppose that X is a normed space and that $x_0 \in S_X$. Then the following are equivalent.*

(a) *The element x_0 is a point of smoothness of B_X.*

(b) *Whenever (x_n^*) is a sequence in S_{X^*} such that $x_n^* x_0 \to 1$, the sequence (x_n^*) is weakly* convergent.*

If x_0 is a point of smoothness of B_X, then the weak limit in* (b) *is the unique member of S_{X^*} that supports B_X at x_0.*

PROOF. Suppose that x_0 is not a point of smoothness of B_X. Let x_1^* and x_2^* be two different members of S_{X^*} that support B_X at x_0. Then the sequence $(x_1^*, x_2^*, x_1^*, x_2^*, \dots)$ is not weakly* convergent, so (b) fails. Therefore (b) $\Rightarrow$ (a).

Suppose conversely that x_0 is a point of smoothness of B_X. Let x_0^* be the element of S_{X^*} that supports B_X at x_0 and let (x_n^*) be a sequence in S_{X^*} such that $x_n^* x_0 \to 1$. Let (x_α^*) be a subnet of (x_n^*). By the Banach-Alaoglu theorem, the net (x_α^*) has a subnet (x_β^*) weakly* convergent to some y^* in B_{X^*}. Now $y^* x_0 = \lim_\beta x_\beta^* x_0 = 1$, so $y^* = x_0^*$. Since every subnet of (x_n^*) has a subnet that is weakly* convergent to x_0^*, it follows that (x_n^*) is weakly* convergent to x_0^*. This proves that (a) $\Rightarrow$ (b) and establishes the final conclusion in the statement of the theorem. ■

As is true for rotundity, smoothness is inherited by subspaces.

5.4.20 Proposition. *If a normed space is smooth, then so is each of its subspaces.*

PROOF. Suppose that M is a subspace of a smooth normed space X, that $x \in S_M$, and that $y \in M$. The left-hand Gateaux derivatives of the norms of M and X at x in the direction y are obviously equal, and the same is true for the corresponding right-hand Gateaux derivatives. It follows immediately that every point of S_M is a point of smoothness of B_M, so M is smooth. ■

Another characteristic that smoothness shares with rotundity is that the existence of the property for a normed space is determined by its presence in the two-dimensional subspaces of the space.

5.4.21 Proposition. *A normed space is smooth if and only if each of its two-dimensional subspaces is smooth.*

PROOF. Suppose that X is a nonsmooth normed space. Then there is an x in S_X and a y in X such that the norm of X is not Gateaux differentiable at x in the direction y. Since X is neither zero- nor one-dimensional, there is a two-dimensional subspace M of X that contains both x and y. The norm of M is not Gateaux differentiable at x in the direction y, so M is not smooth. ■

Though smoothness is inherited by subspaces, it is not always inherited by quotient spaces. S. L. Troyanski gave an example in his 1970 paper [231] of a smooth Banach space with a closed subspace M such that X/M is not smooth.

It is true that smoothness is inherited by direct sums of normed spaces when it is present in each of the summands.

5.4.22 Theorem. *Suppose that $X_1, \dots, X_n$ are normed spaces. Then $X_1 \oplus \cdots \oplus X_n$ is smooth if and only if each X_j is smooth.*

PROOF. It may be assumed that $n = 2$, for then an obvious induction argument proves the general case. Since each of X_1 and X_2 is isometrically isomorphic to a subspace of $X_1 \oplus X_2$, each is smooth when $X_1 \oplus X_2$ is so. Suppose conversely that X_1 and X_2 are smooth. Let (x_1, x_2) be an element of $S_{X_1 \oplus X_2}$ and x^* and y^* elements of $S_{(X_1 \oplus X_2)^*}$ such that $x^*(x_1, x_2) = y^*(x_1, x_2) = 1$. Then there are elements x_1^* and y_1^* of X_1^* and elements x_2^* and y_2^* of X_2^* such that $x^*(z_1, z_2) = x_1^* z_1 + x_2^* z_2$ and $y^*(z_1, z_2) = y_1^* z_1 + y_2^* z_2$ whenever $(z_1, z_2) \in X_1 \oplus X_2$; see Theorem 1.10.13. It follows from Cauchy's inequality for real Euclidean 2-space that

$$\begin{aligned} 1 &= x_1^* x_1 + x_2^* x_2 \\ &\le \|x_1^*\| \, \|x_1\| + \|x_2^*\| \, \|x_2\| \\ &\le \left(\|x_1^*\|^2 + \|x_2^*\|^2\right)^{1/2} \left(\|x_1\|^2 + \|x_2\|^2\right)^{1/2} \\ &= \|x^*\| \, \|(x_1, x_2)\| \\ &= 1, \end{aligned}$$

and, since there must actually be equality throughout, that $\big(\|x_1^*\|, \|x_2^*\|\big)$ is a real multiple of $\big(\|x_1\|, \|x_2\|\big)$. Therefore $\|x_1^*\| = \|x_1\|$ and $\|x_2^*\| = \|x_2\|$. It also follows that $x_1^* x_1 = \|x_1^*\| \, \|x_1\| = \|x_1\|^2$ and $x_2^* x_2 = \|x_2\|^2$. The same argument applied to y^* shows that $\|y_1^*\| = \|x_1\|$, $\|y_2^*\| = \|x_2\|$, $y_1^* x_1 = \|x_1\|^2$, and $y_2^* x_2 = \|x_2\|^2$. However, it is an easy consequence of the smoothness of X_1 that there is a unique member z^* of X_1^* such that $\|z^*\| = \|x_1\|$ and $z^* x_1 = \|x_1\|^2$. Therefore $x_1^* = y_1^*$, and similarly $x_2^* = y_2^*$, so $x^* = y^*$. By Corollary 5.4.3, the normed space $X_1 \oplus X_2$ is smooth. ∎

The rest of this section is devoted to a closer look at the relationship between the points of the unit sphere of a normed space X and the members of S_{X^*} that support B_X at those points, as well as the connection between that relationship and smoothness. In this context it is natural to study the behavior of the following set-valued map.

5.4.23 Definition. Suppose that X is a normed space. For each x in S_X, let

$$\nu(x) = \{ x^* : x^* \in S_{X^*}, \, x^* x = 1 \};$$

that is, let $\nu(x)$ be the subset of S_{X^*} whose members are the support functionals for B_X that support B_X at x. Then ν is the *spherical image map* for S_X.

The convention will be adopted that *with all notation as in the preceding definition, when the map ν is everywhere singleton-valued it will be treated* *in this text as if its values were members of S_{X^*} rather than singleton subsets of S_{X^*}.* The following result, which is essentially just a restatement of Corollary 5.4.3, says that this convention will be invoked precisely when X is smooth.

5.4.24 Proposition. *A normed space is smooth if and only if the spherical image map for the unit sphere of the space is singleton-valued.*

Rotundity can also be characterized in terms of this map. The following is just a restatement of Corollary 5.1.16.

5.4.25 Proposition. *Suppose that X is a normed space and that ν is the spherical image map for S_X. Then X is rotund if and only if $\nu(x_1)$ and $\nu(x_2)$ are disjoint whenever x_1 and x_2 are different members of S_X.*

5.4.26 Corollary. *A normed space is rotund and smooth if and only if the spherical image map for its unit sphere is singleton-valued and one-to-one.*

Spherical image maps for unit spheres of normed spaces are always weakly*-compact-valued.

5.4.27 Proposition. *Suppose that X is a normed space, that ν is the spherical image map for S_X, and that $x \in S_X$. Then $\nu(x)$ is weakly* compact.*

PROOF. Since B_{X^*} is weakly* compact, all that needs to be shown is that $\nu(x)$ is weakly* closed. Suppose that (x_α^*) is a net in $\nu(x)$, that $x^* \in X^*$, and that $x_\alpha^* \xrightarrow{w^*} x^*$. Then $x^* \in B_{X^*}$. Since $x^*x = \lim_\alpha x_\alpha^* x = 1$, it follows that $x^* \in S_{X^*}$ and therefore that $x^* \in \nu(x)$, so $\nu(x)$ is weakly* closed. ∎

When a normed space is smooth and the spherical image map for its unit sphere is treated as point-valued rather than set-valued, it is natural to ask if the map has any continuity properties. More generally, one might ask if the set-valued spherical image map for the unit sphere of an arbitrary normed space has any properties analogous to continuity. In fact, it does.

5.4.28 Theorem. *Suppose that ν is the spherical image map for the unit sphere of a normed space X and that G is a weakly* open subset of X^*. Then $\{x : x \in S_X,\ \nu(x) \subseteq G\}$ is an open subset of S_X.*

PROOF. Suppose to the contrary that $\{x : x \in S_X,\ \nu(x) \subseteq G\}$ is not open in S_X. Then there is an x_0 in S_X and a sequence (x_n) in S_X converging to x_0 such that $\nu(x_0) \subseteq G$ but $\nu(x_n) \nsubseteq G$ whenever $n \in \mathbb{N}$. For each n, let x_n^* be a member of $\nu(x_n) \setminus G$. By the weak* compactness of B_{X^*}, there is some subnet (x_α^*) of (x_n^*) that is weakly* convergent to some x_0^* in B_{X^*}. For each α,

$$|x_0^* x_0 - 1| \leq |x_0^* x_0 - x_\alpha^* x_0| + |x_\alpha^* x_0 - x_\alpha^* x_\alpha| \leq |x_0^* x_0 - x_\alpha^* x_0| + \|x_0 - x_\alpha\|,$$

so passing to the limit shows that $x_0^* x_0 = 1$ and therefore that $x_0^* \in S_{X^*}$. It follows that $x_0^* \in \nu(x_0)$, so x_0^* is in the weakly* open set G even though the net (x_α^*) that is weakly* convergent to x_0^* lies entirely outside G. This contradiction finishes the proof. ■

5.4.29 Corollary. *The spherical image map for the unit sphere of a smooth normed space is norm-to-weak* continuous.*

It follows from this corollary that the spherical image map for the unit sphere of a smooth normed space is norm-to-weak continuous if the space is reflexive, and is even norm-to-norm continuous if the space is finite-dimensional. These stronger continuity conditions can be used to define and characterize strengthened smoothness conditions for normed spaces, as will be seen in Section 5.6.

Exercises

5.37 (a) Give an example of a rotund Banach space that is not smooth. Exercise 5.35 might help.

(b) Give an example of a smooth Banach space that is not rotund.

5.38 This exercise is based on the results of Exercise 5.1. Suppose that (X_n) is a sequence of normed spaces.

(a) Let $\ell_\infty((X_n))$ be the collection of all sequences (x_n) such that $x_n \in X_n$ for each n and $\sup_n \|x_n\|$ is finite. Let $\|(x_n)\|_\infty = \sup_n \|x_n\|$ for each member (x_n) of $\ell_\infty((X_n))$. Show that $\ell_\infty((X_n))$ is a vector space when given the obvious vector space operations and that $\|\cdot\|_\infty$ is a norm on this vector space. Show also that $\ell_\infty((X_n))$ is a Banach space if and only if each X_n is a Banach space.

(b) Suppose that $1 \leq p < \infty$ and that q is conjugate to p. For each member (x_n^*) of $\ell_q((X_n^*))$, let

$$(T(x_n^*))(x_n) = \sum_n x_n^* x_n$$

whenever $(x_n) \in \ell_p((X_n))$. Show that T is an isometric isomorphism from $\ell_q((X_n^*))$ onto $\ell_p((X_n))^*$. A look at the proof of Theorem C.12 could be helpful.

5.39 This exercise is based on Exercise 5.38. Suppose that (X_n) is a sequence of normed spaces and that $1 < p < \infty$. Prove that $\ell_p((X_n))$ is smooth if and only if each (X_n) is smooth.

5.40 Suppose that X is a normed space.

(a) Prove that X^* is rotund if and only if X/M is smooth whenever M is a closed subspace of X such that X/M is two-dimensional.

(b) Prove that X^* is smooth if and only if X/M is rotund whenever M is a closed subspace of X such that X/M is two-dimensional.

(c) Conclude from (a) and (b) that for X to be rotund (respectively, smooth) it is *sufficient* that every two-dimensional quotient space of X be rotund (respectively, smooth). (However, *sufficient* cannot be replaced by *necessary* as can be seen from the examples cited in this section of rotund (respectively, smooth) Banach spaces whose duals are not smooth (respectively, rotund).)

5.41 Let $\ell_{1,r}$ be the nonreflexive rotund Banach space constructed in Example 5.1.8 and studied further in Example 5.1.22. Use Exercise 5.40 to prove that $\ell_{1,r}^*$ is not smooth.

5.42 Suppose that K is a compact Hausdorff space with more than one element. Show that $\text{rca}(K)$ is not smooth.

5.43 This exercise assumes some knowledge of inner product spaces and their duals; see, for example, [24]. Prove that every inner product space is smooth. Exercise 5.4 might help.

5.44 (V. L. Šmulian, 1939 [222]). Suppose that X is a normed space and that $x_0^* \in S_{X^*}$. Show that the following are equivalent.

(a) The element x_0^* is a point of smoothness of B_{X^*}.

(b) Whenever (x_n) is a sequence in S_X such that $x_0^* x_n \to 1$, the sequence (x_n) is weakly Cauchy.

5.45 This exercise depends on Exercise 5.44 and on James's theorem from either of the optional Sections 1.13 and 2.9. Show that a weakly sequentially complete normed space is reflexive if its dual space is smooth. Conclude that there is no norm $\|\cdot\|_{S^*}$ on ℓ_1 equivalent to its standard norm such that $(\ell_1, \|\cdot\|_{S^*})^*$ is smooth.

5.46 (D. F. Cudia, 1964 [44]). Suppose that X is a normed space and that Q is the natural map from X into X^{**}. Then X^* is $Q(X)$-*rotund* if, whenever $x_1^*, x_2^* \in S_{X^*}$ and there is an x^{**} in $S_{Q(X)}$ such that $x^{**}x_1^* = x^{**}x_2^* = 1$, it follows that $x_1^* = x_2^*$. Also, the space X^* is $Q(X)$-*smooth* if, whenever $x_1^{**}, x_2^{**} \in S_{Q(X)}$ and there is an x^* in S_{X^*} such that $x_1^{**}x^* = x_2^{**}x^* = 1$, it follows that $x_1^{**} = x_2^{**}$.

(a) Prove that X is rotund if and only if X^* is $Q(X)$-smooth.

(b) Prove that X is smooth if and only if X^* is $Q(X)$-rotund.

5.47 Characterize the members (α_n) of S_{c_0} that are points of smoothness of B_{c_0} by means of some property of the terms of (α_n). Do the same for the members of S_{ℓ_1} that are points of smoothness of B_{ℓ_1}.

5.48 Prove that if X is a Banach space that is not zero- or one-dimensional, then there is a nonsmooth norm on X equivalent to its original norm.

5.49 Suppose that x_0 is an element of the unit sphere of a normed space X. Prove that x_0 is a point of smoothness of B_X if and only if the map $y \mapsto G_+(x_0, y)$ is a real-linear functional on X.

5.50 (a) This exercise requires James's theorem from either of the optional Sections 1.13 and 2.9. Find a characterization of the reflexive Banach spaces among all Banach spaces in terms of the behavior of the spherical image map for the unit sphere.

(b) Find a characterization of the reflexive, rotund, smooth Banach spaces among all Banach spaces in terms of the behavior of the spherical image map for the unit sphere.

5.51 Suppose that X is a normed space, that Q is the natural map from X into X^{**}, and that ν and ν^* are the spherical image maps for S_X and S_{X^*} respectively.

(a) Prove that X is rotund if and only if $Q^{-1}\big(\nu^*\big(\nu(x)\big)\big) = \{x\}$ whenever $x \in S_X$. (If A is a set and Φ is a set-valued map with domain A, then $\Phi(S) = \bigcup\{\,\Phi(s) : s \in S\,\}$ whenever $S \subseteq A$.)

(b) Prove that X is smooth if and only if $\nu\big(Q^{-1}\big(\nu^*(x^*)\big)\big) = \{x^*\}$ whenever $x^* \in S_{X^*}$.

5.5 Uniform Smoothness

This section and the next are based on strengthenings of the notion of Gateaux differentiability of a norm. The following result is a good starting point for understanding the basis for the strengthenings.

5.5.1 Proposition. *Suppose that X is a normed space. Then the following are equivalent. (The variable t in the expressions is real-valued.)*

(a) *The space X is smooth.*

(b) $\displaystyle\lim_{t\to 0}\frac{\|x+ty\|-\|x\|}{t}$ *exists whenever $x \in S_X$ and $y \in X$.*

(c) $\displaystyle\lim_{t\to 0+}\frac{\frac{1}{2}\big(\|x+ty\|+\|x-ty\|\big)-1}{t} = 0$ *whenever $x \in S_X$ and $y \in X$.*

Now suppose that X is smooth. Then the expression whose limit is being taken in (b) *converges to its limit uniformly for x in S_X when y is a fixed element of S_X, or for y in S_X when x is a fixed element of S_X, or for (x, y) in $S_X \times S_X$, if and only if the expression whose limit is being taken in* (c) *converges to its limit uniformly under the same condition on x and y.*

PROOF. The equivalence of (a) and (b) is just a restatement of the fact that X is smooth if and only if its norm is Gateaux differentiable. Observe next that if $x \in S_X$ and $y \in X$, then

$$\begin{aligned}\frac{G_+(x,y)-G_-(x,y)}{2} &= \frac{1}{2}\lim_{t\to 0+}\left(\frac{\|x+ty\|-\|x\|}{t}-\frac{\|x-ty\|-\|x\|}{-t}\right)\\ &= \lim_{t\to 0+}\frac{\frac{1}{2}\big(\|x+ty\|+\|x-ty\|\big)-1}{t},\end{aligned}$$

from which it follows that the limit in (c) exists, and furthermore that the limit is 0 if and only if the norm of X is Gateaux differentiable at x in the direction y. The equivalence of (a) and (c) follows immediately.

Suppose that X is smooth, that $x, y \in S_X$, and that $t > 0$. It follows from Lemma 5.4.14 that

$$\begin{aligned}&\left|\frac{\|x+ty\|-\|x\|}{t}-G(x,y)\right|+\left|\frac{\|x-ty\|-\|x\|}{-t}-G(x,y)\right|\\ &\qquad=\frac{\|x+ty\|-\|x\|}{t}-G(x,y)+G(x,y)-\frac{\|x-ty\|-\|x\|}{-t}\\ &\qquad=2\,\frac{\frac{1}{2}\big(\|x+ty\|+\|x-ty\|\big)-1}{t}\\ &\qquad=2\left|\frac{\frac{1}{2}\big(\|x+ty\|+\|x-ty\|\big)-1}{t}-0\right|,\end{aligned}$$

from which the claims about uniform convergence follow easily. ■

With all notation as in the statement of the preceding proposition, it is reasonable to say that X is uniformly smooth if the expression whose limit is being taken in (c) converges to 0 uniformly for (x, y) in $S_X \times S_X$; that is, if, for every positive ϵ, there is a positive δ_ϵ depending only on ϵ such that

$$\frac{\frac{1}{2}\big(\|x+ty\|+\|x-ty\|\big)-1}{t}<\epsilon \tag{5.1}$$

whenever $0 < t < \delta_\epsilon$ and $x, y \in S_X$. Notice that the expression on the left side of (5.1) is positive since

$$\tfrac{1}{2}\big(\|x+ty\|+\|x-ty\|\big)-1 \geq \tfrac{1}{2}\|2x\|-1=0,$$

and therefore no absolute value symbols are needed around it.

The actual definition of uniform smoothness to be used in this book will be given in a form equivalent to the one stated in the preceding paragraph. To set the stage for this definition, suppose for the moment that X is a normed space that is not zero-dimensional, and let

$$\rho_X(t)=\sup\{\tfrac{1}{2}\big(\|x+ty\|+\|x-ty\|\big)-1 : x,y\in S_X\}$$

whenever $t > 0$. Then $0 \leq \rho_X(t) \leq t$ for each positive t. It is clear after only a moment's thought that the following are equivalent.

(a) For every positive ϵ there is a positive δ_ϵ such that (5.1) holds whenever $0 < t < \delta_\epsilon$ and $x, y \in S_X$.

(b) For every positive ϵ there is a positive δ_ϵ such that $\rho_X(t)/t < \epsilon$ whenever $0 < t < \delta_\epsilon$.

Consequently, the formal definition of uniform smoothness about to be given is equivalent to the one stated above in terms of uniform convergence.

5.5.2 Definition. Suppose that X is a normed space. Define a function $\rho_X\colon (0,+\infty) \to [0,+\infty)$ by the formula

$$\rho_X(t) = \sup\left\{\tfrac{1}{2}\big(\|x+ty\| + \|x-ty\|\big) - 1 : x, y \in S_X\right\}$$

if $X \neq \{0\}$, and by the formula

$$\rho_X(t) = \begin{cases} 0 & \text{if } 0 < t < 1; \\ t-1 & \text{if } t \geq 1 \end{cases}$$

if $X = \{0\}$. Then ρ_X is the *modulus of smoothness* of X. The space X is *uniformly smooth* if $\lim_{t\to 0^+} \rho_X(t)/t = 0$. The abbreviation US is used for this property.

Here are a few simple properties of the modulus of smoothness that will shed some light on portions of the preceding definition.

5.5.3 Proposition. *Suppose that ρ_X is the modulus of smoothness of a normed space X. Then $\rho_X(t)/t$ is nondecreasing as t increases. Also,*

$$\max\{0, t-1\} \leq \rho_X(t) \leq t$$

whenever $t > 0$.

PROOF. It may be assumed that $X \neq \{0\}$. Suppose that $0 < t_1 < t_2$ and that $x, y \in S_X$. By Lemma 5.4.14,

$$\begin{aligned}
\frac{\frac{1}{2}\big(\|x+t_1y\| + \|x-t_1y\|\big) - 1}{t_1} &= \frac{1}{2}\left(\frac{\|x+t_1y\| - \|x\|}{t_1} - \frac{\|x-t_1y\| - \|x\|}{-t_1}\right) \\
&\leq \frac{1}{2}\left(\frac{\|x+t_2y\| - \|x\|}{t_2} - \frac{\|x-t_2y\| - \|x\|}{-t_2}\right) \\
&= \frac{\frac{1}{2}\big(\|x+t_2y\| + \|x-t_2y\|\big) - 1}{t_2},
\end{aligned}$$

from which it follows that $\rho_X(t_1)/t_1 \leq \rho_X(t_2)/t_2$.

Now fix a positive t. Since $0 \leq \rho_X(t) \leq t$, all that remains to be proved is that $\rho_X(t) \geq t - 1$. However, this is an immediate consequence of the fact that if $x, y \in S_X$, then

$$\tfrac{1}{2}\big(\|x+ty\| + \|x-ty\|\big) - 1 \geq \tfrac{1}{2}\|2ty\| - 1 = t - 1,$$

which finishes the proof. ■

With all notation as in the statement of the preceding proposition, it follows that $\rho_X(t)/t$ converges to some nonnegative real number as $t \to 0^+$, so the content of the requirement that $\lim_{t\to 0^+} \rho_X(t)/t = 0$ for X to be uniformly smooth is that the limit be 0; existence of the limit is not an issue. Also, it follows from Proposition 5.5.3 that the modulus of smoothness of a zero-dimensional normed space has been defined to be as small as it can be while still satisfying the inequalities of that proposition.

An approach to uniform smoothness only slightly different from the one taken above is based on condition (b) of Proposition 5.5.1 and involves the last differentiability criterion in the following definition. The other forms of differentiability in this definition will not be used until Section 5.6 but are included here for comparison.

5.5.4 Definition. Suppose that X is a normed space. Then the norm of X is

(a) *uniformly Gateaux differentiable* if

$$\lim_{t\to 0} \frac{\|x+ty\| - \|x\|}{t}$$

exists for every x in S_X and y in X, where the variable t is real-valued, and furthermore the convergence is uniform for x in S_X whenever y is a fixed element of S_X;

(b) *Fréchet differentiable* if the limit in (a) exists for every x in S_X and y in X and the convergence is uniform for y in S_X whenever x is a fixed element of S_X;

(c) *uniformly Fréchet differentiable* if the limit in (a) exists for every x in S_X and y in X and the convergence is uniform for (x, y) in $S_X \times S_X$.

5.5.5 Definition. A normed space is *uniformly Fréchet smooth* if its norm is uniformly Fréchet differentiable. The symbol UF is used for this property.

The following result is an immediate consequence of the equivalence of (b) and (c) in Proposition 5.5.1, the observations about uniform convergence in that proposition, and the discussion preceding Definition 5.5.2.

5.5.6 Proposition. *A normed space is uniformly smooth if and only if it is uniformly Fréchet smooth. In symbols,* (US) $\Leftrightarrow$ (UF).

Of course, every uniformly Fréchet differentiable norm is Gateaux differentiable, which proves the most basic property of uniformly smooth normed spaces.

5.5.7 Proposition. *Every uniformly smooth normed space is smooth.*

Another very basic property of uniform smoothness follows immediately from its definition.

5.5.8 Proposition. *Every normed space that is isometrically isomorphic to a uniformly smooth normed space is itself uniformly smooth.*

As one might suspect, the preceding result does not generalize to isomorphisms. For example, the space ℓ_1^2 is not uniformly smooth since it is not even smooth, but it is isomorphic to ℓ_2^2 which will be shown to be uniformly smooth in Corollary 5.5.17.

Uniform smoothness implies stronger continuity properties for the spherical image map than does smoothness. In fact, uniformly smooth normed spaces can be characterized among all smooth normed spaces in terms of the continuity of this map, as is apparent from the following lemma. This lemma will have a further application in Section 5.6.

5.5.9 Lemma. *Suppose that X is a smooth normed space and that ν is the spherical image map for S_X. Then ν, treated as a point-valued map, is norm-to-norm continuous if and only if the norm of X is Fréchet differentiable, and is uniformly norm-to-norm continuous on S_X if and only if the norm of X is uniformly Fréchet differentiable.*

PROOF. It follows from the smoothness of X that if $x \in S_X$, then the Gateaux derivative $G(x,y)$ of the norm at x in the direction y exists whenever $y \in X$, and the map $y \mapsto G(x,y)$ is $\operatorname{Re}\nu(x)$.

Suppose first that the norm of X is uniformly Fréchet differentiable but that ν is not uniformly norm-to-norm continuous on S_X. This implies the existence of a positive s and sequences (z_n) and (x_n) in S_X such that $\|z_n - x_n\| \to 0$ but $\|\nu(z_n) - \nu(x_n)\| > s$ for each n. It follows that there is a sequence (y_n) in S_X such that $\operatorname{Re}\big(\nu(z_n) - \nu(x_n)\big)(y_n) > s$ for each n. By the uniform Fréchet differentiability of the norm of X, there is a positive t such that whenever $n \in \mathbb{N}$,

$$\left|\frac{\|x_n + ty_n\| - \|x_n\|}{t} - \operatorname{Re}\big(\nu(x_n)\big)(y_n)\right| = \left|\frac{\|x_n + ty_n\| - \|x_n\|}{t} - G(x_n, y_n)\right| < \frac{s}{2}$$

and therefore

$$\|x_n + ty_n\| - \|x_n\| - \operatorname{Re}\big(\nu(x_n)\big)(ty_n) < \frac{ts}{2}.$$

For each positive integer n,

$$\begin{aligned}\operatorname{Re}\big(\nu(z_n)\big)(x_n + ty_n - z_n) &= \operatorname{Re}\big(\nu(z_n)\big)(x_n + ty_n) - 1\\ &\le \|x_n + ty_n\| - \|x_n\|,\end{aligned}$$

which implies that

$$\mathrm{Re}\big(\nu(z_n)\big)(ty_n) \le \|x_n + ty_n\| - \|x_n\| + \mathrm{Re}\big(\nu(z_n)\big)(z_n - x_n)$$

and therefore that

$$\begin{aligned} ts &< \mathrm{Re}\big(\nu(z_n) - \nu(x_n)\big)(ty_n) \\ &\le \|x_n + ty_n\| - \|x_n\| - \mathrm{Re}\big(\nu(x_n)\big)(ty_n) + \mathrm{Re}\big(\nu(z_n)\big)(z_n - x_n) \\ &< \frac{ts}{2} + \|z_n - x_n\|. \end{aligned}$$

It follows that $\|z_n - x_n\|$ does not tend to 0, a contradiction. Therefore the uniform Fréchet differentiability of the norm of X implies the uniform norm-to-norm continuity of ν on S_X. A similar argument, with the sequence (x_n) replaced by a single element x of S_X, shows that ν is norm-to-norm continuous whenever the norm of X is Fréchet differentiable.

Now suppose conversely that ν is uniformly norm-to-norm continuous on S_X. Fix a positive ϵ. Then there is a δ_ϵ such that $0 < \delta_\epsilon \le 1$ for which $\big\|\nu\big(\|x + ty\|^{-1}(x + ty)\big) - \nu(x)\big\| < \epsilon$ when $x, y \in S_X$ and $0 < t < \delta_\epsilon$; notice that $x + ty \ne 0$ since $0 < t < 1$. If $x, y \in S_X$ and $0 < t < \delta_\epsilon$, then

$$\begin{aligned} \mathrm{Re}\big(\nu(x)\big)(ty) &= \mathrm{Re}\big(\nu(x)\big)(x + ty) - 1 \\ &\le \|x + ty\| - \|x\| \\ &\le \|x + ty\| \ \mathrm{Re}\Big(\nu\big(\|x + ty\|^{-1}(x + ty)\big)\Big) \\ &\qquad\qquad \big(\|x + ty\|^{-1}(x + ty) - \|x + ty\|^{-1}x\big) \\ &= \mathrm{Re}\Big(\nu\big(\|x + ty\|^{-1}(x + ty)\big)\Big)(ty), \end{aligned}$$

which in turn implies that

$$\begin{aligned} 0 &\le \|x + ty\| - \|x\| - \mathrm{Re}\big(\nu(x)\big)(ty) \\ &\le \mathrm{Re}\Big(\nu\big(\|x + ty\|^{-1}(x + ty)\big)\Big)(ty) - \mathrm{Re}\big(\nu(x)\big)(ty) \\ &\le t\big\|\nu\big(\|x + ty\|^{-1}(x + ty)\big) - \nu(x)\big\| < t\epsilon \end{aligned}$$

and therefore that

$$\left|\frac{\|x + ty\| - \|x\|}{t} - G(x, y)\right| = \left|\frac{\|x + ty\| - \|x\|}{t} - \mathrm{Re}\big(\nu(x)\big)(y)\right| < \epsilon.$$

It follows that the norm of X is uniformly Fréchet differentiable. Fixing x but otherwise applying the same argument also proves that the norm of X is Fréchet differentiable whenever ν is norm-to-norm continuous. ■

The preceding lemma and the equivalence of uniform Fréchet smoothness and uniform smoothness immediately produce the following result.

5.5.10 Theorem. *Suppose that X is a normed space and that ν is the spherical image map for S_X. Then X is uniformly smooth if and only if ν is singleton-valued and, when viewed as a point-valued map, is uniformly norm-to-norm continuous on S_X.*

It was seen in Section 5.4 that if the dual space of a normed space X is either rotund or smooth, then X has the other property, but the presence of one of the properties in X does not imply that of the other in X^*. It turns out that there is a more nearly complete duality between uniform rotundity and uniform smoothness, since a normed space has either property if and only if its dual space has the other. Early proofs of this or of results that readily imply this can be found in [50] and [224]. The proof to be given here is from [157] and is based on an interesting relationship between the moduli of rotundity and smoothness of a normed space and of its dual.

5.5.11 Lemma. (J. Lindenstrauss, 1963 [154]). *Suppose that δ_X and ρ_X are, respectively, the modulus of rotundity and modulus of smoothness of a normed space X and that δ_{X^*} and ρ_{X^*} are the corresponding moduli of X^*. Then*

$$\rho_{X^*}(t) = \sup\left\{ \frac{t\epsilon}{2} - \delta_X(\epsilon) : 0 \le \epsilon \le 2 \right\}$$

and

$$\rho_X(t) = \sup\left\{ \frac{t\epsilon}{2} - \delta_{X^*}(\epsilon) : 0 \le \epsilon \le 2 \right\}$$

when $t > 0$.

PROOF. It is easy to check that the formulas work if $X = \{0\}$, so it will be assumed that $X \neq \{0\}$. For each positive t,

$$\begin{aligned} 2\rho_{X^*}(t) &= \sup\{ \|x^* + ty^*\| + \|x^* - ty^*\| - 2 : x^*, y^* \in S_{X^*} \} \\ &= \sup\{ \operatorname{Re} x^*x + t \operatorname{Re} y^*x + \operatorname{Re} x^*y - t \operatorname{Re} y^*y - 2 : \\ &\qquad\qquad x, y \in S_X,\ x^*, y^* \in S_{X^*} \} \\ &= \sup\{ \|x + y\| + t\|x - y\| - 2 : x, y \in S_X \} \\ &= \sup\{ \|x + y\| + t\epsilon - 2 : x, y \in S_X,\ \|x - y\| \ge \epsilon,\ 0 \le \epsilon \le 2 \} \\ &= \sup\Big\{ t\epsilon - 2\inf\big\{ 1 - \big\|\tfrac{1}{2}(x + y)\big\| : x, y \in S_X,\ \|x - y\| \ge \epsilon \big\} : \\ &\qquad\qquad 0 \le \epsilon \le 2 \Big\} \\ &= \sup\{ t\epsilon - 2\delta_X(\epsilon) : 0 \le \epsilon \le 2 \}. \end{aligned}$$

Dividing through by 2 gives the first of the two claimed formulas. The other is proved exactly the same way after exchanging the roles of X and X^*. ■

Šmulian's 1940 version of the following result is actually that uniform Fréchet differentiability of the norm is the condition dual to uniform rotundity.

5.5.12 Theorem. (V. L. Šmulian, 1940 [224]). *A normed space is uniformly rotund if and only if its dual space is uniformly smooth, and is uniformly smooth if and only if its dual space is uniformly rotund.*

PROOF. Let X be a normed space and let δ_X and ρ_{X^*} be, respectively, the modulus of rotundity of X and the modulus of smoothness of X^*. By the preceding lemma,

$$\frac{\rho_{X^*}(t)}{t} = \sup\left\{\frac{\epsilon}{2} - \frac{\delta_X(\epsilon)}{t} : 0 \le \epsilon \le 2\right\}$$

when $t > 0$. Suppose first that X is uniformly rotund, and let s be a positive number. To prove that X^* is uniformly smooth, it is enough to find a positive t_s such that $\rho_{X^*}(t_s)/t_s \le s$, because $\rho_{X^*}(t)/t$ is nonnegative for each positive t and is nonincreasing as $t \to 0^+$. To this end, let $\epsilon_s = \min\{2, 2s\}$ and let $t_s = \delta_X(\epsilon_s)$. If $0 \le \epsilon \le \epsilon_s$, then

$$\frac{\epsilon}{2} - \frac{\delta_X(\epsilon)}{t_s} \le \frac{\epsilon}{2} \le \frac{\epsilon_s}{2} \le s,$$

while if $\epsilon_s < \epsilon \le 2$, then the fact that δ_X is nondecreasing assures that

$$\frac{\epsilon}{2} - \frac{\delta_X(\epsilon)}{t_s} \le 1 - \frac{\delta_X(\epsilon_s)}{t_s} = 0 < s.$$

It follows that $\rho_{X^*}(t_s)/t_s \le s$, so X^* is uniformly smooth. An analogous argument with the roles of X and X^* exchanged proves that X is uniformly smooth when X^* is uniformly rotund.

Now suppose that X^* is uniformly smooth and that $0 < \epsilon \le 2$. Let the positive number t_ϵ be such that $\rho_{X^*}(t_\epsilon)/t_\epsilon \le \epsilon/4$. It follows that

$$\frac{\epsilon}{2} - \frac{\delta_X(\epsilon)}{t_\epsilon} \le \frac{\rho_{X^*}(t_\epsilon)}{t_\epsilon} \le \frac{\epsilon}{4},$$

so $\delta_X(\epsilon) \ge t_\epsilon\epsilon/4 > 0$. Therefore X is uniformly rotund. The argument just given, with the roles of X and X^* exchanged, then proves that X^* is uniformly rotund when X is uniformly smooth. ■

It follows from the preceding result and the Milman-Pettis theorem that if X is a uniformly smooth normed space, then X^* is uniformly rotund and therefore reflexive, which in turn implies the reflexivity of X provided that X is complete.

5.5.13 Theorem. (V. L. Šmulian, 1940 [224]). *Every uniformly smooth Banach space is reflexive.*

It is finally time to give some examples.

5.5.14 Example. Since every zero- or one-dimensional normed space has a uniformly rotund dual space, each such space is uniformly smooth. The uniform smoothness of zero-dimensional normed spaces also follows trivially from the definition of the property.

5.5.15 Example. The Banach space $(c_0, \|\cdot\|_b)$ of Example 5.4.13 is smooth but not reflexive, and therefore is an example of a smooth Banach space that is not uniformly smooth.

The main examples of uniformly smooth Banach spaces arise from the following theorem.

5.5.16 Theorem. *Suppose that μ is a positive measure on a σ-algebra Σ of subsets of a set Ω and that $1 < p < \infty$. Then $L_p(\Omega, \Sigma, \mu)$ is uniformly smooth.*

PROOF. Let q be such that $p^{-1} + q^{-1} = 1$. Then $\big(L_p(\Omega, \Sigma, \mu)\big)^*$ is isometrically isomorphic to the uniformly rotund normed space $L_q(\Omega, \Sigma, \mu)$ and is therefore itself uniformly rotund, so $L_p(\Omega, \Sigma, \mu)$ is uniformly smooth. ■

5.5.17 Corollary. *Suppose that $1 < p < \infty$. Then ℓ_p is uniformly smooth, as is ℓ_p^n whenever n is a nonnegative integer.*

It is probably becoming obvious that much of the theory of uniform smoothness can be obtained from the theory of uniform rotundity and the duality between the two properties. This lode will be mined several more times in the rest of this section. Here is another application of the method.

5.5.18 Proposition. *If a normed space is finite-dimensional, then it is uniformly smooth if and only if it is smooth.*

PROOF. Suppose that X is a finite-dimensional smooth normed space. Then X^* is finite-dimensional and rotund, and so is uniformly rotund, which implies that X is uniformly smooth. ■

As with smoothness, uniform smoothness is inherited by subspaces. The reason for this is that thinning a normed space to a subspace cannot increase the modulus of smoothness.

5.5.19 Lemma. *Suppose that M is a subspace of a normed space X and that ρ_M and ρ_X are the respective moduli of smoothness of the two spaces. Then $\rho_M(t) \le \rho_X(t)$ when $t > 0$.*

PROOF. This follows directly from the definition of ρ_M, with a little help from Proposition 5.5.3 when $M = \{0\}$. ■

5.5.20 Proposition. *Every subspace of a uniformly smooth normed space is uniformly smooth.*

Unlike smoothness but like uniform rotundity, uniform smoothness is always inherited by quotient spaces.

5.5.21 Lemma. *Suppose that M is a closed subspace of a normed space X and that $\rho_{X/M}$ and ρ_X are the moduli of smoothness of X/M and X respectively. Then $\rho_{X/M}(t) \le \rho_X(t)$ when $t > 0$.*

PROOF. Fix a positive t and let $\delta_{M^\perp}$ and δ_{X^*} be the respective moduli of rotundity of $M^\perp$ and X^*. Since $\delta_{M^\perp}(\epsilon) \ge \delta_{X^*}(\epsilon)$ when $0 \le \epsilon \le 2$ and since $(X/M)^*$ is isometrically isomorphic to $M^\perp$, it follows from Lemma 5.5.11 that

$$\begin{aligned}\rho_{X/M}(t) &= \sup\left\{\frac{t\epsilon}{2} - \delta_{M^\perp}(\epsilon) : 0 \le \epsilon \le 2\right\}\\ &\le \sup\left\{\frac{t\epsilon}{2} - \delta_{X^*}(\epsilon) : 0 \le \epsilon \le 2\right\}\\ &= \rho_X(t),\end{aligned}$$

as claimed. ■

5.5.22 Theorem. *If M is a closed subspace of a uniformly smooth normed space X, then X/M is uniformly smooth.*

Uniform smoothness is also inherited by direct sums from the summands.

5.5.23 Theorem. *Suppose that $X_1, \dots, X_n$ are normed spaces. Then $X_1 \oplus \cdots \oplus X_n$ is uniformly smooth if and only if each X_j is uniformly smooth.*

PROOF. If $X_1 \oplus \cdots \oplus X_n$ is uniformly smooth, then each X_j is uniformly smooth since it is isometrically isomorphic to a subspace of $X_1 \oplus \cdots \oplus X_n$. Suppose conversely that each X_j is uniformly smooth. By Theorem 5.2.25, the normed space $X_1^* \oplus \cdots \oplus X_n^*$ is uniformly rotund because each of its summands is uniformly rotund, so $(X_1 \oplus \cdots \oplus X_n)^*$ is uniformly rotund since it is isometrically isomorphic to $X_1^* \oplus \cdots \oplus X_n^*$. Therefore $X_1 \oplus \cdots \oplus X_n$ is uniformly smooth. ■

It is apparent from the results of this section that there are many analogies between the behavior of uniformly smooth normed spaces and that of uniformly rotund normed spaces. In fact, there is a much stronger connection between the two properties than their duality and these analogies would suggest. It turns out that a normed space has a uniformly smooth norm equivalent to its original norm if and only if it has a uniformly rotund norm equivalent to its original norm; that is, that $\langle \mathrm{US} \rangle \Leftrightarrow \langle \mathrm{UR} \rangle$; and therefore that the $\langle \mathrm{US} \rangle$ normed spaces are exactly the superreflexive ones. A proof and references can be found in [56, pp. 169–173].

Exercises

5.52 Recall that a real-valued function f defined on a possibly unbounded subinterval I of $\mathbb{R}$ is *convex* if

$$f\big(st_1 + (1-s)t_2\big) \le sf(t_1) + (1-s)f(t_2)$$

whenever $t_1, t_2 \in I$ and $0 < s < 1$. Prove that for every normed space X, the modulus of smoothness of X is a convex function.

5.53 Prove that for every normed space X, the modulus of smoothness of X is a continuous function. (It is not at all necessary to base the argument on the result of Exercise 5.52, but feel free to use that result if you have proved it.)

5.54 The modulus of smoothness of a normed space X is sometimes defined by the formula

$$\rho_X(t) = \sup\{\tfrac{1}{2}\big(\|x+y\| + \|x-y\|\big) - 1 : x \in B_X,\ y \in X,\ \|y\| \le t\}$$

when $X \neq \{0\}$. Show that this definition is equivalent to the one given in this section for normed spaces that are not zero-dimensional.

5.55 Suppose that X is a normed space. Prove that the following are equivalent.

(a) The space X is uniformly smooth.

(b) For each positive η there is a positive ϵ_η such that if $x \in S_X$, $y \in X$, and $\|x-y\| \le \epsilon_\eta$, then $\|x+y\| \ge \|x\| + \|y\| - \eta\|x-y\|$.

The condition given in (b) is sometimes used as the definition of uniform smoothness. See, for example, [56, p. 147].

5.56 Suppose that X is a normed space that is neither zero- nor one-dimensional. Let ρ_M represent the modulus of smoothness of M whenever M is a subspace of X. For each positive t, let

$$f(t) = \sup\{\rho_M(t) : M \text{ is a two-dimensional subspace of } X\}.$$

Prove that X is uniformly smooth if and only if $f(t)/t \to 0$ as $t \to 0^+$.

5.57 (M. M. Day, 1944 [50]). Prove that a normed space X is uniformly rotund if and only if

$$\inf\{\delta_{X/M}(\epsilon) : M \text{ is a closed subspace of } X \text{ and } X/M \text{ is two-dimensional}\} > 0$$

whenever $0 < \epsilon \le 2$, where $\delta_{X/M}$ is the modulus of rotundity of the quotient space X/M. Exercise 5.56 could be helpful.

5.58 Prove that every uniformly rotund normed space has a uniformly rotund completion and that every uniformly smooth normed space has a uniformly smooth completion.

5.59 Find an explicit formula, in terms only of functions one might encounter in a precalculus course, for the modulus of smoothness of Euclidean 2-space. The result of Exercise 5.15 could help, but it is not crucial.

5.60 This exercise assumes some knowledge of inner product spaces and their duals; see, for example, [24]. Prove that every inner product space is uniformly smooth. Exercise 5.16 might help.

5.61 Prove that, for every normed space, the modulus of smoothness of the space is the same as the modulus of smoothness of its second dual.

5.6 Generalizations of Uniform Smoothness

The purpose of this section is to examine several conditions that lie between smoothness and uniform smoothness, primarily by seeing what happens when the norm is Gateaux differentiable in some uniform sense that is not quite so strong as the uniform Fréchet sense characterizing uniform smoothness. For example, requiring the norm to be only Fréchet differentiable as in Definition 5.5.4 (b) results in the following generalization of uniform smoothness.

5.6.1 Definition. A normed space is *Fréchet smooth* if its norm is Fréchet differentiable. The symbol F is used for this property.

It is clear that every uniformly Fréchet differentiable norm is Fréchet differentiable and that every Fréchet differentiable norm is Gateaux differentiable. This can be restated in the language of smoothness as follows.

5.6.2 Proposition. *Every uniformly smooth normed space is Fréchet smooth, and every Fréchet smooth normed space is smooth. In symbols,* (US) $\Rightarrow$ (F) $\Rightarrow$ (S).

As has already been proved in Lemma 5.5.9, Fréchet smoothness can be characterized in terms of the continuity of the spherical image map for the unit sphere. The following result is just a restatement of the pertinent portion of that lemma.

5.6.3 Theorem. *Suppose that X is a normed space and that ν is the spherical image map for S_X. Then X is Fréchet smooth if and only if ν is singleton-valued and, when viewed as a point-valued map, is norm-to-norm continuous.*

Since smoothness of the dual space of a normed space X implies that X is rotund, and since uniform smoothness of X^* implies that X is uniformly rotund, one would expect that Fréchet smoothness of X^* would imply some rotundity condition for X lying between simple rotundity and uniform rotundity. In fact, it turns out that Fréchet smoothness of X^* is *equivalent* to X being strongly rotund, a fact due to Šmulian. The proof to be given here will closely follow Šmulian's, and is based on four lemmas concerning Cauchy and weakly* Cauchy norming sequences.

5.6.4 Definition. Suppose that X is a normed space and that Z is a subspace of X^* such that

$$\|x\| = \sup\{\, |z^*x| : z^* \in S_Z\}$$

whenever $x \in X$. Then Z is a *norming subspace* for X in X^*. If $x \in X$ and (z_n^*) is a sequence in S_Z such that $z_n^*x \to \|x\|$, then (z_n^*) is a *norming sequence* for x in S_Z.

5.6.5 Lemma. (V. L. Šmulian, 1939 [221]). *Suppose that X is a normed space, that $x \in S_X$, that Z is a norming subspace for X in X^*, that (z_n^*) is a weakly* Cauchy norming sequence for x in S_Z, that $y \in S_X$, and that t is a nonzero real number such that $|t| \le 1/4$. Then there exists a sequence $\big(y_n^*(\,\cdot\,; y, t)\big)$ in S_Z, where the notation indicates that the linear functionals y_n^* depend on y and t, such that*

(1) $\lim_n \operatorname{Re} y_n^*(x; y, t)$ *and* $\lim_n \operatorname{Re} y_n^*(y; y, t)$ *both exist;*

(2) $\left|\dfrac{\|x+ty\| - \|x\|}{t} - \lim_n \operatorname{Re} z_n^* y\right| \le |\lim_n \operatorname{Re} z_n^* y - \lim_n \operatorname{Re} y_n^*(y; y, t)|;$ *and*

(3) $|\lim_n \operatorname{Re} y_n^*(x; y, t) - 1| \le 2|t|$.

PROOF. Let $\big(y_n^*(\,\cdot\,; y, t)\big)$ be a sequence in S_Z such that

$$\|x+ty\| = \lim_n y_n^*(x+ty; y, t) = \lim_n\big(\operatorname{Re} y_n^*(x; y, t) + t\operatorname{Re} y_n^*(y; y, t)\big).$$

By thinning the sequence $\big(y_n^*(\,\cdot\,; y, t)\big)$ if necessary, it may be assumed that $\lim_n \operatorname{Re} y_n^*(x; y, t)$ and $\lim_n \operatorname{Re} y_n^*(y; y, t)$ both exist. Then

$$\begin{aligned}\lim_n \operatorname{Re} y_n^*(x; y, t) &= \|x+ty\| - t\lim_n \operatorname{Re} y_n^*(y; y, t)\\ &\ge \|x\| - |t|\,\|y\| - |t|\,|\lim_n \operatorname{Re} y_n^*(y; y, t)|\\ &\ge 1 - 2|t|,\end{aligned}$$

so

$$0 \le 1 - \lim_n \operatorname{Re} y_n^*(x; y, t) \le 2|t|,$$

which proves that (3) holds. Notice next that

$$\begin{aligned}\|x+ty\| - \|x\| &\ge |\lim_n \operatorname{Re} z_n^*(x+ty)| - 1\\ &= (1 + t\lim_n \operatorname{Re} z_n^* y) - 1\\ &= t\lim_n \operatorname{Re} z_n^* y.\end{aligned}$$

Also,

$$\begin{aligned}\|x+ty\| - \|x\| &= \lim_n \operatorname{Re} y_n^*(x;y,t) + t\lim_n \operatorname{Re} y_n^*(y;y,t) - 1\\ &\le t\lim_n \operatorname{Re} y_n^*(y;y,t).\end{aligned}$$

Therefore

$$0 \le \|x+ty\| - \|x\| - t\lim_n \operatorname{Re} z_n^* y \le t\lim_n \operatorname{Re} y_n^*(y;y,t) - t\lim_n \operatorname{Re} z_n^* y.$$

Taking absolute values and dividing through by $|t|$ proves that (2) holds. ■

5.6.6 Lemma. (V. L. Šmulian, 1939 [221]). *Suppose that X is a normed space, that Z is a norming subspace for X in X^*, and that a sequence in S_Z is Cauchy if there is an element of S_X for which the sequence is a norming sequence. Then X is Fréchet smooth.*

PROOF. Throughout this proof, references to (1), (2), and (3) are to the conditions in the conclusion of Lemma 5.6.5. Fix an element x of S_X and a norming sequence (z_n^*) for x in S_Z. For each y in S_X and each nonzero real number t such that $|t| \le 1/4$, let $\big(y_n^*(\,\cdot\,;y,t)\big)$ be a sequence in S_Z satisfying (1), (2), and (3). A moment's thought about (2) and the definition of a Fréchet differentiable norm shows that it suffices to prove that

$$|\lim_n \operatorname{Re} z_n^* y - \lim_n \operatorname{Re} y_n^*(y;y,t)|$$

converges to 0 as $t \to 0$ uniformly for y in S_X. Suppose to the contrary that this does not happen. Then there must be a sequence (t_m) of nonzero real numbers such that $|t_m| \le 1/4$ for each m and $\lim_m t_m = 0$, a sequence (y_m) in S_X, and a positive ϵ such that

$$|\lim_n \operatorname{Re} z_n^* y_m - \lim_n \operatorname{Re} y_n^*(y_m;y_m,t_m)| \ge \epsilon \tag{5.2}$$

for every m. Let (j_m) be an increasing sequence of positive integers such that

$$|\operatorname{Re} y_{j_m}^*(x;y_m,t_m) - \lim_n \operatorname{Re} y_n^*(x;y_m,t_m)| < \frac{1}{m} \tag{5.3}$$

and

$$|\operatorname{Re} y_{j_m}^*(y_m;y_m,t_m) - \lim_n \operatorname{Re} y_n^*(y_m;y_m,t_m)| < \frac{1}{m} \tag{5.4}$$

for each m. Then $\operatorname{Re} y_{j_m}^*(x;y_m,t_m) \to 1$ by (3) and the fact that (5.3) holds for each m, so $\big(y_{j_m}^*(\,\cdot\,;y_m,t_m)\big)$ is a norming sequence for x in S_Z. Since (5.2) and (5.4) hold for each m, it also follows that

$$|\lim_n \operatorname{Re} z_n^* y_m - \operatorname{Re} y_{j_m}^*(y_m;y_m,t_m)| > \frac{\epsilon}{2}$$

for all sufficiently large m, which together with the fact that (z_n^*) is Cauchy implies that

$$\|z_n^* - y_{j_m}^*(\,\cdot\,; y_m, t_m)\| \ge |\operatorname{Re} z_n^* y_m - \operatorname{Re} y_{j_m}^*(y_m; y_m, t_m)| > \frac{\epsilon}{4}$$

whenever n and m are greater than some positive integer n_0. However, this cannot be, since the sequence

$$\big(z_1^*,\, y_{j_1}^*(\,\cdot\,; y_1, t_1),\, z_2^*,\, y_{j_2}^*(\,\cdot\,; y_2, t_2),\, z_3^*,\, y_{j_3}^*(\,\cdot\,; y_3, t_3),\, \dots \big)$$

is a norming sequence for x in S_Z and therefore is itself Cauchy. ■

5.6.7 Lemma. (V. L. Šmulian, 1940 [224]). *Suppose that X is a normed space, that $x \in S_X$, that Z is a norming subspace for X in X^*, and that (z_n^*) is a weakly* Cauchy norming sequence for x in S_Z. Then*

$$\begin{aligned}(\operatorname{Re} z^* y)&(\operatorname{Re} z^* x) - \lim_n \operatorname{Re} z_n^* y \\ &\le \big(1 - |\operatorname{Re} z^* x|\big)\left(\frac{1}{t} + 2\right) + \frac{\|x + t[y - (\lim_n \operatorname{Re} z_n^* y)x]\| - \|x\|}{t}\end{aligned}$$

whenever $z^* \in S_Z$, $y \in S_X$, and $t > 0$.

PROOF. For each real number s, let

$$\operatorname{sign} s = \begin{cases} +1 & \text{if } s \ge 0; \\ -1 & \text{if } s < 0. \end{cases}$$

Fix a z^* in S_Z, a y in S_X, and a positive t. Then

$$\begin{aligned}(\operatorname{Re} z^* y)(\operatorname{Re} z^* x) &- \lim_n \operatorname{Re} z_n^* y \\ &= (\operatorname{Re} z^* y)(\operatorname{sign} \operatorname{Re} z^* x) - |\operatorname{Re} z^* x| \lim_n \operatorname{Re} z_n^* y \\ &\quad + (\operatorname{Re} z^* y)(\operatorname{Re} z^* x - \operatorname{sign} \operatorname{Re} z^* x) + \big(|\operatorname{Re} z^* x| - 1\big) \lim_n \operatorname{Re} z_n^* y \\ &\le (\operatorname{Re} z^* y)(\operatorname{sign} \operatorname{Re} z^* x) - |\operatorname{Re} z^* x| \lim_n \operatorname{Re} z_n^* y \\ &\quad + |\operatorname{Re} z^* x - \operatorname{sign} \operatorname{Re} z^* x| + \big(1 - |\operatorname{Re} z^* x|\big) \\ &= [\operatorname{Re} z^* y - (\operatorname{Re} z^* x) \lim_n \operatorname{Re} z_n^* y] \operatorname{sign} \operatorname{Re} z^* x + 2\big(1 - |\operatorname{Re} z^* x|\big) \\ &= 2\big(1 - |\operatorname{Re} z^* x|\big) + \frac{1 - |\operatorname{Re} z^* x|}{t} \\ &\quad + \frac{\big(\operatorname{Re} z^* x + t[\operatorname{Re} z^* y - (\operatorname{Re} z^* x) \lim_n \operatorname{Re} z_n^* y]\big) \operatorname{sign} \operatorname{Re} z^* x - \|x\|}{t} \\ &\le \big(1 - |\operatorname{Re} z^* x|\big)\left(\frac{1}{t} + 2\right) + \frac{\|x + t[y - (\lim_n \operatorname{Re} z_n^* y)x]\| - \|x\|}{t},\end{aligned}$$

as claimed. ■

5.6.8 Lemma. (V. L. Šmulian, 1940 [224]). *Suppose that X is a Fréchet smooth normed space, that $x \in S_X$, that Z is a norming subspace for X in X^*, and that (z_n^*) is a norming sequence for x in S_Z. Then (z_n^*) is Cauchy.*

PROOF. Let x^* be the unique element of S_{X^*} that supports B_X at x. Then $z_n^* \xrightarrow{w^*} x^*$ by Theorem 5.4.19. By Theorem 5.4.17,

$$\lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t} = \operatorname{Re} x^* y$$

whenever $y \in X$. Therefore

$$\lim_{t \to 0} \frac{\|x + t[y - (\operatorname{Re} x^* y)x]\| - \|x\|}{t} = 0$$

whenever $y \in S_X$. Fix a positive ϵ and let δ_ϵ be a positive real number such that

$$\frac{\|x + t[y - (\operatorname{Re} x^* y)x]\| - \|x\|}{t} \le \frac{\epsilon}{4}$$

whenever $0 < t \le \delta_\epsilon$ and $y \in S_X$; the Fréchet differentiability of the norm of X guarantees the existence of such a δ_ϵ. It follows from Lemma 5.6.7 that

$$(\operatorname{Re} z_n^* y)(\operatorname{Re} z_n^* x) - \operatorname{Re} x^* y \le (1 - |\operatorname{Re} z_n^* x|)\left(\frac{1}{\delta_\epsilon} + 2\right) + \frac{\epsilon}{4}$$

for each y in S_X and each positive integer n. Since $-y \in S_X$ whenever $y \in S_X$, it even follows that

$$|(\operatorname{Re} z_n^* y)(\operatorname{Re} z_n^* x) - \operatorname{Re} x^* y| \le (1 - |\operatorname{Re} z_n^* x|)\left(\frac{1}{\delta_\epsilon} + 2\right) + \frac{\epsilon}{4}$$

for each y in S_X and each n. It is an easy consequence of this and the convergence of $\operatorname{Re} z_n^* x$ to 1 that there is a positive integer n_0 such that

$$|\operatorname{Re} z_{n_1}^* y - \operatorname{Re} z_{n_2}^* y| \le \epsilon$$

whenever $y \in S_X$ and $n_1, n_2 \ge n_0$, and therefore that

$$\|z_{n_1}^* - z_{n_2}^*\| = \sup\{\, |\operatorname{Re} z_{n_1}^* y - \operatorname{Re} z_{n_2}^* y| : y \in S_X \} \le \epsilon$$

whenever $n_1, n_2 \ge n_0$. The sequence (z_n^*) is therefore Cauchy. ■

5.6.9 Theorem. (V. L. Šmulian, 1940 [224]). *A normed space is strongly rotund if and only if its dual space is Fréchet smooth.*

PROOF. Let X be a normed space and let Q be the natural map from X into X^{**}. It may be assumed that $X \ne \{0\}$. Then $Q(X)$ is a norming subspace for X^* in X^{**}. By Theorem 5.3.17, the space X is strongly rotund if and only if the following condition holds: Whenever $x^* \in S_{X^*}$ and (x_n) is a sequence in X such that (Qx_n) is a norming sequence for x^* in $S_{Q(X)}$, the sequence (Qx_n) is Cauchy. By Lemmas 5.6.6 and 5.6.8, this last condition is equivalent to X^* being Fréchet smooth. ■

Since Fréchet smoothness is obviously preserved by isometric isomorphisms, a reflexive normed space is Fréchet smooth if and only if its second dual is Fréchet smooth, which together with the preceding theorem produces the following result.

5.6.10 Corollary. *A reflexive normed space is Fréchet smooth if and only if its dual space is strongly rotund.*

As will be seen in a moment, this corollary does not in general extend to nonreflexive Banach spaces.

Since the dual space of a normed space X is a norming subspace for X when $X \neq \{0\}$, the following result is another immediate consequence of Lemmas 5.6.6 and 5.6.8 when $X \neq \{0\}$, and is trivially true when $X = \{0\}$. Notice that the Cauchy property of the lemmas has been replaced by convergence since X^* is complete.

5.6.11 Theorem. (V. L. Šmulian, 1940 [224]). *Suppose that X is a normed space. Then the following are equivalent.*

(a) *The space X is Fréchet smooth.*

(b) *Whenever (x_n^*) is a sequence in S_{X^*} for which there is an element x of S_X such that $\operatorname{Re} x_n^* x \to 1$, the sequence (x_n^*) converges.*

By Theorem 5.3.17, the dual space of a normed space X is strongly rotund if and only if the following holds: Whenever (x_n^*) is a sequence in S_{X^*} for which there is an element x^{**} of $S_{X^{**}}$ such that $\operatorname{Re} x^{**} x_n^* \to 1$, the sequence (x_n^*) converges. Condition (b) of Theorem 5.6.11 is analogous, except that the only members of $S_{X^{**}}$ considered are those that come from $Q(X)$, where Q is the natural map from X into X^{**}. In fact, this condition (b) is properly weaker than the strong rotundity of X^*, since there are Fréchet smooth Banach spaces whose duals are not strongly rotund. One example of this follows from the fact that there is a Fréchet smooth norm $\|\cdot\|_F$ for c_0 equivalent to the usual norm for c_0; see [56, p. 160]. As was mentioned in Section 5.3 at the end of the discussion of strong rotundity, it can be shown that every strongly rotund Banach space is reflexive, so $(c_0, \|\cdot\|_F)^*$ cannot be strongly rotund.

The following, at least, does hold, by an easy argument based on the observations about sequential characterizations given at the beginning of the preceding paragraph.

5.6.12 Theorem. *If the dual space of a normed space X is strongly rotund, then X is Fréchet smooth.*

The example $(c_0, \|\cdot\|_F)$ given above can also be used to show that Fréchet smoothness is different from uniform smoothness, since the Banach space $(c_0, \|\cdot\|_F)$ is Fréchet smooth but cannot be uniformly smooth because it is not reflexive. In fact, the three properties of uniform smoothness, Fréchet smoothness, and smoothness are distinct even in the presence of reflexivity, as can be shown by applying Theorem 5.6.9 to some examples by Mark Smith from [219]. Among Smith's examples are two reflexive normed spaces $(\ell_2, \|\cdot\|_L)$ and $(\ell_2, \|\cdot\|_W)$, each formed by giving ℓ_2 a norm equivalent to its original norm, such that $(\ell_2, \|\cdot\|_L)$ is (K) but not (UR) and $(\ell_2, \|\cdot\|_W)$ is (R) but not (K). It follows that $(\ell_2, \|\cdot\|_L)^*$ is (F) but not (US), while $(\ell_2, \|\cdot\|_W)^*$ is (S) but not (F).

It is time to move on to the other major generalization of uniform smoothness to be examined in this section.

5.6.13 Definition. A normed space is *uniformly Gateaux smooth* if its norm is uniformly Gateaux differentiable. The symbol UG is used for this property.

It is clear from the definitions that every uniformly Fréchet differentiable norm is uniformly Gateaux differentiable and that every uniformly Gateaux differentiable norm is Gateaux differentiable, which translates into the language of smoothness as follows.

5.6.14 Proposition. *Every uniformly smooth normed space is uniformly Gateaux smooth, and every uniformly Gateaux smooth normed space is smooth. In symbols,* (US) $\Rightarrow$ (UG) $\Rightarrow$ (S).

The main fact to be proved here about uniform Gateaux smoothness is a characterization of the property in terms of a rotundity property of the dual space. The proof is, essentially, Šmulian's original one.

5.6.15 Theorem. (V. L. Šmulian, 1940 [224]). *A normed space is uniformly Gateaux smooth if and only if its dual space is weakly* uniformly rotund.*

PROOF. Let X be a normed space, which may be assumed not to be $\{0\}$. Suppose first that X is uniformly Gateaux smooth, that ϵ is a fixed real number such that $0 < \epsilon \le 2$, and that y is a fixed element of S_X. Let x be an arbitrary element of S_X, let z^* be an arbitrary element of S_{X^*}, and let x^* be the unique element of S_{X^*} that supports B_X at x. Since X^* is a norming subspace for X in X^*, it follows from Lemma 5.6.7 that whenever

$t > 0$,

$$(\operatorname{Re} z^*y)(\operatorname{Re} z^*x) - \operatorname{Re} x^*y \le \left(1 - |\operatorname{Re} z^*x|\right)\left(\frac{1}{t} + 2\right) + \frac{\|x + t[y - (\operatorname{Re} x^*y)x]\| - \|x\|}{t}.$$

By Theorem 5.4.17,

$$\lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t} = \operatorname{Re} x^*y,$$

so the uniform Gateaux differentiability of the norm of X implies the existence of a $\theta_{\epsilon,y}$, not dependent on the member x of S_X, such that $0 < \theta_{\epsilon,y} < 1$ and

$$\begin{aligned}\frac{\|x + t[y - (\operatorname{Re} x^*y)x]\| - \|x\|}{t} &= \frac{\|(1 - t\operatorname{Re} x^*y)x + ty\| - \|(1 - t\operatorname{Re} x^*y)x\|}{t} - \operatorname{Re} x^*y \\ &= \frac{\|x + (1 - t\operatorname{Re} x^*y)^{-1}ty\| - \|x\|}{(1 - t\operatorname{Re} x^*y)^{-1}t} - \operatorname{Re} x^*y \\ &\le \frac{\epsilon}{16}\end{aligned}$$

whenever $0 < t \le \theta_{\epsilon,y}$. Thus,

$$(\operatorname{Re} z^*y)(\operatorname{Re} z^*x) - \operatorname{Re} x^*y \le \left(1 - |\operatorname{Re} z^*x|\right)\left(\frac{1}{\theta_{\epsilon,y}} + 2\right) + \frac{\epsilon}{16}.$$

Replacing x and x^* by their negatives in this inequality then shows that

$$|(\operatorname{Re} z^*y)(\operatorname{Re} z^*x) - \operatorname{Re} x^*y| \le \left(1 - |\operatorname{Re} z^*x|\right)\left(\frac{1}{\theta_{\epsilon,y}} + 2\right) + \frac{\epsilon}{16}.$$

It is worth pointing out again that $\theta_{\epsilon,y}$ does not depend on the member x of S_X or the member z^* of S_{X^*}. Therefore

$$|(\operatorname{Re} z_1^*y)(\operatorname{Re} z_1^*x) - (\operatorname{Re} z_2^*y)(\operatorname{Re} z_2^*x)| \le \left(2 - |\operatorname{Re} z_1^*x| - |\operatorname{Re} z_2^*x|\right)\left(\frac{1}{\theta_{\epsilon,y}} + 2\right) + \frac{\epsilon}{8}$$

whenever $z_1^*, z_2^* \in S_{X^*}$ and $x \in S_X$. It follows that there is a $\delta_{\epsilon,y}$ such that $0 < \delta_{\epsilon,y} < 1/4$ and

$$|(\operatorname{Re} z_1^*y)(\operatorname{Re} z_1^*x) - (\operatorname{Re} z_2^*y)(\operatorname{Re} z_2^*x)| \le \frac{\epsilon}{4}$$

whenever $z_1^*, z_2^* \in S_{X^*}$, $x \in S_X$, and $\operatorname{Re} z_1^* x, \operatorname{Re} z_2^* x \ge 1 - 4\delta_{\epsilon,y}$. After decreasing $\delta_{\epsilon,y}$ a bit if necessary, it follows that

$$|\operatorname{Re} z_1^* y - \operatorname{Re} z_2^* y| < \frac{\epsilon}{2}$$

whenever $z_1^*, z_2^* \in S_{X^*}$, $x \in S_X$, and $\operatorname{Re} z_1^* x, \operatorname{Re} z_2^* x \ge 1 - 4\delta_{\epsilon,y}$.

For the moment, fix two elements z_1^* and z_2^* of S_{X^*} such that

$$\left\|\tfrac{1}{2}(z_1^* + z_2^*)\right\| \ge 1 - \delta_{\epsilon,y}.$$

Then there is an x in S_X such that $\frac{1}{2}(\operatorname{Re} z_1^* x + \operatorname{Re} z_2^* x) \ge 1 - 2\delta_{\epsilon,y}$, which in turn implies that $\operatorname{Re} z_1^* x, \operatorname{Re} z_2^* x \ge 1 - 4\delta_{\epsilon,y}$ and thus that

$$|\operatorname{Re} z_1^* y - \operatorname{Re} z_2^* y| < \frac{\epsilon}{2}.$$

If $\mathbb{F} = \mathbb{C}$, then $\left\|\frac{1}{2}(iz_1^* + iz_2^*)\right\| \ge 1 - \delta_{\epsilon,y}$, from which it follows that

$$|\operatorname{Im} z_1^* y - \operatorname{Im} z_2^* y| = |\operatorname{Re}(iz_2^* y) - \operatorname{Re}(iz_1^* y)| < \frac{\epsilon}{2}.$$

In any case, it follows that $|z_1^* y - z_2^* y| < \epsilon$ whenever $z_1^*, z_2^* \in S_{X^*}$ and $\left\|\frac{1}{2}(z_1^* + z_2^*)\right\| \ge 1 - \delta_{\epsilon,y}$. Let δ_{X^*} be the w*UR modulus of X^*. Then

$$\begin{aligned}\delta_{X^*}(\epsilon, y) &= \inf\left\{ 1 - \left\|\tfrac{1}{2}(z_1^* + z_2^*)\right\| : z_1^*, z_2^* \in S_{X^*}, |z_1^* y - z_2^* y| \ge \epsilon \right\} \\ &\ge \delta_{\epsilon,y} \\ &> 0,\end{aligned}$$

and so X^* is weakly* uniformly rotund.

Now suppose conversely that X^* is weakly* uniformly rotund. Then X^* is rotund, so X is smooth. As in the first part of this proof, let y be a fixed element of S_X. Let x be an arbitrary member of S_X, let t be an arbitrary real number such that $0 < |t| \le 1/4$, and let x_x^* be the unique element of S_{X^*} that supports B_X at x. By Lemma 5.6.5, there exists a sequence $(y_n^*(\,\cdot\,; y, x, t))$ in S_{X^*} such that

(4) $\lim_n \operatorname{Re} y_n^*(x; y, x, t)$ and $\lim_n \operatorname{Re} y_n^*(y; y, x, t)$ both exist;

(5) $\left|\dfrac{\|x + ty\| - \|x\|}{t} - \operatorname{Re} x_x^* y\right| \le |\operatorname{Re} x_x^* y - \lim_n \operatorname{Re} y_n^*(y; y, x, t)|$; and

(6) $|\lim_n \operatorname{Re} y_n^*(x; y, x, t) - 1| \le 2|t|$.

By (5), the theorem will be proved once it is shown that

$$|\operatorname{Re} x_x^* y - \lim_n \operatorname{Re} y_n^*(y; y, x, t)|$$

converges to 0 as $t \to 0$ uniformly for x in S_X. Suppose to the contrary that this does not occur. Then there is a sequence (t_m) of nonzero real numbers such that $|t_m| \le 1/4$ for each m and $\lim_m t_m = 0$, a sequence (x_m) in S_X and a corresponding sequence (x_m^*) in S_{X^*} such that each x_m^* supports B_X at x_m, and a positive ϵ such that

$$|\operatorname{Re} x_m^* y - \lim_n \operatorname{Re} y_n^*(y; y, x_m, t_m)| \ge \epsilon \tag{5.5}$$

for each m. Let (j_m) be an increasing sequence of positive integers such that

$$|\operatorname{Re} y_{j_m}^*(y; y, x_m, t_m) - \lim_n \operatorname{Re} y_n^*(y; y, x_m, t_m)| < \frac{1}{m} \tag{5.6}$$

and

$$|\operatorname{Re} y_{j_m}^*(x_m; y, x_m, t_m) - \lim_n \operatorname{Re} y_n^*(x_m; y, x_m, t_m)| < \frac{1}{m} \tag{5.7}$$

for each m. Since (5.5) and (5.6) hold for each m, it follows that

$$|x_m^* y - y_{j_m}^*(y; y, x_m, t_m)| \ge |\operatorname{Re} x_m^* y - \operatorname{Re} y_{j_m}^*(y; y, x_m, t_m)| > \frac{\epsilon}{2}$$

whenever $m \ge 2/\epsilon$. Consequently,

$$\left\|\tfrac{1}{2}\left(x_m^* + y_{j_m}^*(\,\cdot\,; y, x_m, t_m)\right)\right\| \le 1 - \delta_{X^*}(\epsilon/2, y) < 1$$

whenever $m \ge 2/\epsilon$. However, it also follows from (6) that

$$\lim_m\left(\lim_n \operatorname{Re} y_n^*(x_m; y, x_m, t_m)\right) = 1,$$

which together with the fact that (5.7) holds for every m implies that

$$\operatorname{Re} y_{j_m}^*(x_m; y, x_m, t_m) > 1 - \delta_{X^*}(\epsilon/2, y)$$

whenever m is sufficiently large. Therefore for large values of m,

$$\begin{aligned}
\left\|\tfrac{1}{2}\left(x_m^* + y_{j_m}^*(\,\cdot\,; y, x_m, t_m)\right)\right\| &\ge \tfrac{1}{2}\left(\operatorname{Re} x_m^* x_m + \operatorname{Re} y_{j_m}^*(x_m; y, x_m, t_m)\right) \\
&> \tfrac{1}{2}\left(1 + 1 - \delta_{X^*}(\epsilon/2, y)\right) \\
&> 1 - \delta_{X^*}(\epsilon/2, y) \\
&\ge \left\|\tfrac{1}{2}\left(x_m^* + y_{j_m}^*(\,\cdot\,; y, x_m, t_m)\right)\right\|,
\end{aligned}$$

which is the contradiction needed to finish the proof. ■

It takes only a few moments' reflection on the wUR and w*UR moduli of the dual space of a reflexive normed space X to see that X^* is weakly uniformly rotund if and only if it is weakly* uniformly rotund. The following result is therefore an immediate consequence of Theorem 5.6.15.

5.6.16 Corollary. *A reflexive normed space is uniformly Gateaux smooth if and only if its dual space is weakly uniformly rotund.*

5.6.17 Corollary. *A reflexive normed space is weakly uniformly rotund if and only if its dual space is uniformly Gateaux smooth.*

PROOF. Let X be a reflexive normed space. Then X^* is uniformly Gateaux smooth if and only if X^{**} is weakly uniformly rotund, which happens if and only if X is weakly uniformly rotund since weak uniform rotundity is preserved by isometric isomorphisms. ■

One implication in the preceding corollary survives even in the absence of reflexivity.

5.6.18 Corollary. *If the dual space of a normed space X is uniformly Gateaux smooth, then X is weakly uniformly rotund.*

PROOF. Since X^* is (UG), the space X^{**} is (w*UR) by Theorem 5.6.15. An obvious argument based on the way that X naturally embeds in X^{**} and on the forms of the wUR modulus of X and the w*UR modulus of X^{**} then shows that X is (wUR). ■

Corollary 5.6.17, along with the two examples by Mark Smith mentioned in the discussion following Theorem 5.6.12, can be used to show that uniform Gateaux smoothness is genuinely different from both smoothness and uniform smoothness, even for reflexive normed spaces. Smith showed in [219] that $(\ell_2, \|\cdot\|_W)$ is (wUR) but not (UR) and that $(\ell_2, \|\cdot\|_L)$ is (R) but not (wUR). It follows that $(\ell_2, \|\cdot\|_W)^*$ is (UG) but not (US), while $(\ell_2, \|\cdot\|_L)^*$ is (S) but not (UG). As was previously mentioned, the space $(\ell_2, \|\cdot\|_L)^*$ is (F) while $(\ell_2, \|\cdot\|_W)^*$ is not, so this also shows that neither of the properties of Fréchet smoothness and uniform Gateaux smoothness implies the other.

One further generalization of uniform smoothness deserves mention before leaving the subject. It has been shown in Proposition 5.4.24, Corollary 5.4.29, and Theorems 5.6.3 and 5.5.10 that a normed space X is

(a) smooth,

(b) Fréchet smooth, or

(c) uniformly smooth

if and only if the spherical image map for S_X is singleton-valued and, when viewed as a point-valued map, is

(a) norm-to-weak* continuous,

(b) norm-to-norm continuous, or

(c) uniformly norm-to-norm continuous,

respectively. The following definition fills the one conspicuous gap in this list of continuity conditions.

5.6.19 Definition. A normed space is *very smooth* if the spherical image map for its unit sphere is singleton-valued and, when viewed as a point-valued map, is norm-to-weak continuous. The symbol VS is used for this property.

The final result of this section then follows immediately from the observations made just before the preceding definition.

5.6.20 Proposition. *If a normed space is Fréchet smooth then it is very smooth, and if it is very smooth then it is smooth. In symbols,* (F) $\Rightarrow$ (VS) $\Rightarrow$ (S).

Several conditions equivalent to the property of being very smooth can be inferred from the results in J. R. Giles's article [87], which is a good starting point for the reader interested in this condition.

Exercises

5.62 (A. R. Lovaglia, 1955 [159]). Prove that a normed space is Fréchet smooth if its dual space is locally uniformly rotund.

5.63 (V. L. Šmulian, 1940 [224]). Show that a Banach space is reflexive if its dual space is both strongly rotund and Fréchet smooth. Do not use James's theorem or any result based on it, such as Exercise 5.33, in your argument. (In fact, it follows from Exercise 5.33 and Theorem 5.6.9 that a Banach space is reflexive if its dual space is *either* strongly rotund *or* Fréchet smooth.)

5.64 (V. L. Šmulian, 1940 [224]). Suppose that X is a weakly sequentially complete normed space with Schur's property (as is the case for every normed space isomorphic to ℓ_1; see Example 2.5.24). Prove that if X^* is smooth, then X^* is Fréchet smooth.

In particular, it follows from this exercise that if there is a norm $\|\cdot\|_{S^*}$ on ℓ_1 equivalent to the standard norm of ℓ_1 such that $(\ell_1, \|\cdot\|_{S^*})^*$ is smooth, then $(\ell_1, \|\cdot\|_{S^*})$ is strongly rotund. However, this is a contradiction since it would imply that ℓ_1 is reflexive; see Exercise 5.33. See also Exercise 5.45. As was mentioned in Section 5.4, M. M. Day showed in a 1955 paper [51] that ℓ_∞ cannot be equivalently renormed to be smooth, which is a stronger statement than saying that ℓ_1 cannot be equivalently renormed to make its dual space smooth.

5.65 (a) Suppose that X is a strongly rotund Banach space, that Q is the natural map from X into X^{**}, that $x^{**} \in S_{X^{**}}$, that $x^* \in S_{X^*}$, and that $x^{**}x^* = 1$. Prove that there is an x in S_X such that $Qx = x^{**}$. Do not use any material from any optional sections of this book in your argument.

(b) Use an argument based on the Bishop-Phelps subreflexivity theorem from optional Section 2.11 to prove that every strongly rotund Banach space is reflexive.

5.66 (L. P. Vlasov, 1972, 1973 [236, 238]). A normed space X has the *semi-Radon-Riesz property* if it satisfies the following condition: Whenever (x_n), x, and (x_n^*) are, respectively, a sequence in S_X, an element of S_X, and a sequence in S_{X^*} such that $x_n^*x_n = 1$ for each n and $x_n \xrightarrow{w} x$, it follows that $x_n^*x \to 1$. (This condition was previously encountered in Exercise 2.65, but that exercise is not needed for this one.) Show that every uniformly Gateaux smooth normed space has the semi-Radon-Riesz property.

5.67 (J. R. Giles, 1975 [87]). Suppose that X is a normed space, that Q is the natural map from X into X^{**}, and that every point of $S_{Q(X)}$ is a point of smoothness of $B_{X^{**}}$. Prove that X is very smooth.

Appendix A
Prerequisites

This book is intended primarily as a text for Banach space courses taught to graduate students in mathematics. The following are the prerequisites for a course taught from this book at that level.

1. *A first course in linear algebra that covers bases, linear transformations, dual vector spaces, and related topics in the theory of real and complex vector spaces.* The course should not treat just the finite-dimensional theory.

2. *An acquaintance with elementary properties of the complex plane.* To understand all of the examples, some knowledge is needed of analytic functions and integrals of complex-valued functions of a real variable. However, the actual theory presented in this book does not require properties of complex numbers much beyond elementary facts about moduli and conjugates and the fact that the complex numbers form a complete field.

3. *Courses covering elementary measure theory and the Lebesgue spaces $L_p(\Omega, \Sigma, \mu)$, where $1 \le p \le \infty$ and μ is a nonnegative-extended-real-valued measure on a σ-algebra Σ of subsets of the set Ω.* The material covered should include the completeness of Lebesgue spaces as well as the duality between L_p and L_q when $1 \le p < \infty$, the "exponents" p and q are conjugate, and the measure space is σ-finite. The reader should also have seen the Riesz representation theorem for bounded linear functionals on the real and complex Banach spaces $C(K)$, where K is a compact Hausdorff space. However, the Riesz repre-

sentation theorem is actually needed only to understand examples involving the dual space of $C(K)$, not for the main development of the theory.

4. *A first course in general topology that includes product topologies, metric topologies, the separation axioms, and related topics.* Much of the topology needed for this course is presented at the beginning of Section 2.1, but in a condensed fashion that is probably not ideal for someone seeing the material for the first time.

A course in functional analysis is *not* a prerequisite. The results from functional analysis required for this book, such as the Hahn-Banach theorems, open mapping theorem, closed graph theorem, and uniform boundedness principle, are stated and proved as they are needed.

Chapter 1 can be used as the basis for an undergraduate Banach space course having the following reduced set of prerequisites.

1. *The linear algebra course mentioned above.*

2. *A first course in real analysis.* This course should cover the topology of the real line, continuous real-valued functions on the real line, properties of the Riemann integral, and convergence of sequences and series of real numbers and real-valued functions on the real line.

3. *An introduction to metric spaces.* This introduction should cover most of the material outlined in Appendix B.

The following is a section-by-section description of the changes that should be made to the presentation in Chapter 1 for such a course. Notice that after Section 1.1, most of the changes are modifications or omissions of examples, exercises, and comments that depend on material not covered by the above three prerequisites.

Section 1.1. If complex vector spaces are not to be treated, then use the symbol $\mathbb{F}$ to represent only the field $\mathbb{R}$ and replace the paragraph on page 2 that begins with "It is worth emphasizing ... " by "Since all vector spaces will have $\mathbb{R}$ as their scalar field, the terms *vector space* and *real vector space* will be treated as being synonymous. Similarly, the adjective *real* is to be considered to be redundant whenever it is used to describe some special type of vector space, such as the normed spaces and Banach spaces to be defined in the next section." In the definitions of closure and interior, replace "topological space" by "metric space." In the definitions of closed convex hull and closed linear hull, replace "topology" by "metric." After covering the definition of a metric space, review the properties of metric spaces listed in Appendix B, perhaps treating some of the less-standard ones as exercises. Notice in particular that the term "topology" is defined for metric spaces in Definition B.5, allowing its use in Chapter 1 in the context of metric spaces. Omit Exercise 1.12. If complex vector spaces are not to be treated, then omit Exercises 1.9 and 1.10 and the first part of Exercise 1.7.

Section 1.2. If complex vector spaces are not to be treated, then omit the reference to $\mathbb{C}^n$ in Example 1.2.5. Omit Example 1.2.6, and replace the paragraph following it by the single sentence "The roles of subscripts and superscripts are often exchanged in the notations for the normed spaces of the next three examples." Cover Appendix C through Definition C.9 before covering Examples 1.2.7–1.2.9, and omit the paragraph following Example 1.2.7 as well as the last two sentences of Example 1.2.9. In Example 1.2.10, replace "Hausdorff space" by "metric space" and omit the reference to $L_p[0,1]$. Omit Example 1.2.11, and in the paragraph following it replace "topological space" by "metric space" and omit everything after the argument that $C(K)$ is complete. Cover Theorem C.10 and Corollary C.11 after covering Definition 1.2.12. Omit Example 1.2.14. Delete the reference to the Banach spaces $L_p[0,1]$ in the last sentence before the exercises. Omit Exercises 1.22, 1.23, and 1.28.

Section 1.3. In Corollary 1.3.4, replace "topological space" by "metric space." Omit the paragraph immediately preceding Theorem 1.3.10. In Exercise 1.30, replace "Hausdorff space" by "metric space."

Section 1.4. If complex vector spaces are not to be treated, then omit part (c) of Exercise 1.48.

Section 1.5. Delete the last sentence of the second paragraph of the section. In the third paragraph of the section, leave "topological meaning" as it is, since this reference makes sense in the context of the topology of a metric space, but change "topological space" to "metric space." In the collection of definitions following this paragraph, change the first occurrence of "topological space" to "metric space" and change the last sentence to "In particular, the set A is of the *first* or *second category in itself* if A has that category as a subset of the metric space (A, d_A), where d_A is the restriction of the metric of X to $A \times A$." In Proposition 1.5.3, change "topological space" to "metric space." Omit the first sentence of the last paragraph before the exercises. In Exercises 1.53–1.57, replace "topological space" or "topological spaces" by "metric space" or "metric spaces" respectively. Omit Exercises 1.58, 1.62, and 1.63.

Section 1.6. In Definition 1.6.4, replace "topological space" by "metric space." In the paragraph following that definition, omit all but the last sentence. In the paragraph preceding Corollary 1.6.6, replace "topological space" by "metric space." In Exercise 1.68, replace "topological space" and "Hausdorff space" by "metric space." Omit Exercise 1.70. In Exercise 1.71, replace "Hausdorff space" by "metric space." In Exercise 1.74, replace "topological space" by "metric space."

Section 1.7. Omit Example 1.7.2 and the last sentence of the paragraph preceding it. In the paragraph following the proof of Theorem 1.7.4, omit the three references to a topology. In the last sentence of the paragraph preceding Definition 1.7.10, delete everything following "a *quotient map.*"

Omit the first two paragraphs following the proof of Proposition 1.7.12. If it cannot be assumed that the student has seen the isomorphism theorems for groups, omit the paragraph preceding Theorem 1.7.14 and skip Exercise 1.87.

Section 1.8. Omit Exercise 1.92.

Section 1.9. If complex scalars are not to be treated, then omit the following: beginning with the second sentence of the paragraph preceding Definition 1.9.2, all material through Proposition 1.9.3; the last paragraph of the proof of Theorem 1.9.6; the occurrence of "real-" in the paragraph preceding Definition 1.9.13; the two occurrences of "Re" in the statement of Proposition 1.9.15; the first sentence of the proof of Proposition 1.9.15; and the single occurrence of "Re" in Exercise 1.109 as well as its two occurrences in Exercise 1.110.

Section 1.10. Omit Example 1.10.2 and the two paragraphs following it. Cover Theorem C.12 before covering Example 1.10.3, and replace the first three sentences of that example with the following: "Suppose that $1 \le p < \infty$ and that q is conjugate to p. Then ℓ_p^* can be identified with ℓ_q in a natural way; see Theorem C.12." Omit the last two sentences of the first paragraph of Example 1.10.4. Cover Theorem C.13 before covering Example 1.10.5, then replace the entire text of that example with the following: "Suppose that n is a positive integer, that $1 \le p \le \infty$, and that q is conjugate to p. Then $(\ell_p^n)^*$ can be identified with ℓ_q^n in a natural way; see Theorem C.13." Omit Example 1.10.6. Omit the last two sentences of the paragraph following Theorem 1.10.7. If complex scalars are not to be treated, then omit Exercise 1.113. Omit Exercise 1.121.

Section 1.11. Replace the statement of Theorem 1.11.10 by that of Theorem C.14, and the proof of Theorem 1.11.10 by the statement "See Theorem C.14." Omit the paragraph following the proof of Theorem 1.11.10. Omit Examples 1.11.24 and 1.11.25. Omit Exercises 1.124–1.126, 1.129, and 1.130.

Section 1.12. In the introductory paragraph of the section and in the statement and proof of Proposition 1.12.9, replace each occurrence of "topological space" by "metric space." If complex scalars are not to be treated, then replace the third sentence of the proof of Proposition 1.12.1 by "Let $\mathbf{Q}_0$ be the rationals." Omit Examples 1.12.3–1.12.5. In the paragraph following the proof of Corollary 1.12.12, omit the references to $L_1[0,1]$, $C[0,1]$, $L_\infty[0,1]$, and rca$[0,1]$. Omit Exercises 1.142, 1.143, and 1.145.

Section 1.13. Omit the first sentence of the paragraph following Definition 1.13.2. Omit Example 1.13.7 and Exercise 1.149. If complex scalars are not to be treated, then omit the following: all occurrences of the symbol "Re"; the second paragraph of the section; Proposition 1.13.1; the first sentence of the proof of Theorem 1.13.4; the third sentence of the second paragraph following the proof of Theorem 1.13.4; the proof of Theorem 1.13.15; and Exercise 1.146.

Appendix B
Metric Spaces

The only results from topology that are used in a crucial way in Chapter 1 are about metric spaces. These results are standard fare in most general topology texts, but a development of metric spaces independent of general topology, such as can be found in Kaplansky's *Set Theory and Metric Spaces* [129] and in many elementary real analysis texts, is sufficient. For reference, here are the definitions and results from metric space theory needed for Chapter 1.

B.1 Definition. A *metric space* is a set M with a *metric* or *distance function* $d\colon M \times M \to \mathbb{R}$ such that the following three conditions are satisfied by all x, y, and z in M:

(1) $d(x,y) \geq 0$, and $d(x,y) = 0$ if and only if $x = y$;
(2) $d(x,y) = d(y,x)$;
(3) $d(x,z) \leq d(x,y) + d(y,z)$ (the triangle inequality).

For the rest of this appendix, let $M, M_1, \ldots, M_N$ be sets with respective metrics $d, d_1, \ldots, d_N$.

B.2 Proposition. *If S is a subset of M, then S is itself a metric space with the metric inherited from M.*

B.3 Proposition. *If $x, y, z \in M$, then $\left|d(x,z) - d(y,z)\right| \leq d(x,y)$.*

B.4 Definition. If $x \in M$ and $r > 0$, then the *open ball* in M with center x and radius r is $\{\, y : y \in M,\ d(x,y) < r \,\}$. The corresponding *closed ball* is $\{\, y : y \in M,\ d(x,y) \leq r \,\}$.

B.5 Definition. A subset S of M is *open* if, for every x in S, there is an entire open ball centered at x included in S. The *topology* induced by the metric d is the collection of all open subsets of M.

B.6 Proposition. *The following subsets of M are open.*

(1) *The empty set.*
(2) *The entire set M.*
(3) *Every open ball.*
(4) *Every intersection of finitely many open sets.*
(5) *Every union of open sets, whether or not there are finitely many.*

B.7 Definition. A *neighborhood* of an element x of M is an open set containing x.

B.8 Proposition. *Let S be a subset of M. Then the following are equivalent.*

(a) *The set S is open.*
(b) *The set S can be written as the union of a (possibly empty) collection of open balls.*
(c) *For each x in S, there is an entire neighborhood of x included in S.*

B.9 Definition. A subset S of M is *closed* if its complement is open.

B.10 Proposition. *The following subsets of M are closed.*

(1) *The empty set.*
(2) *The entire set M.*
(3) *Every closed ball.*
(4) *Every finite set.*
(5) *Every intersection of closed sets, whether or not there are finitely many.*
(6) *Every union of finitely many closed sets.*

B.11 Definition. Let (x_n) be a sequence in M. Then (x_n) is *Cauchy* if, for every positive ϵ, there is a positive integer N_ϵ such that $d(x_n, x_m) < \epsilon$ whenever $n, m \geq N_\epsilon$. The sequence (x_n) *converges* to an element x of M, and x is called the *limit* of (x_n), if, for every positive ϵ, there is a positive integer N_ϵ such that $d(x_n, x) < \epsilon$ whenever $n \geq N_\epsilon$. If (x_n) converges to x, then this is denoted by writing $x_n \to x$ or $\lim_n x_n = x$.

B.12 Proposition. *A sequence (x_n) in M converges to an element x of M if and only if the following is true: For every neighborhood U of x, there is a positive integer N_U such that $x_n \in U$ whenever $n \geq N_U$.*

B.13 Proposition. *No sequence in M has more than one limit.*

B.14 Proposition. *If a sequence in M converges to a limit, then every subsequence of that sequence converges to that same limit.*

B.15 Proposition. *Every convergent sequence in M is Cauchy.*

B.16 Proposition. *If a Cauchy sequence in M has a convergent subsequence, then the entire sequence converges to the limit of the subsequence.*

B.17 Definition. A subset of M is *bounded* if it is either empty or included in some open ball.

In the preceding definition, there is no need to make a special case of the empty set unless M is itself empty.

B.18 Proposition. *Every Cauchy sequence in M (and hence every convergent sequence in M) is bounded.*

B.19 Proposition. *If $x_n \to x$ and $y_n \to y$ in M, then $d(x_n, y_n) \to d(x, y)$.*

B.20 Proposition. *A subset S of M is closed if and only if every sequence in S that converges in M has its limit in S.*

B.21 Definition. The *closure* of a subset S of M, denoted by $\overline{S}$, is the smallest closed set that includes S, that is, the intersection of all closed sets that include S. The *interior* of S, denoted by S°, is the largest open subset of S, that is, the union of all open subsets of S. The *boundary* of S, denoted by ∂S, is the set $\overline{S} \cap \overline{M \setminus S}$.

B.22 Proposition. *Let S be a subset of M and x an element of M. Then the following are equivalent.*

(a) $x \in \overline{S}$.

(b) *Every neighborhood of x intersects S.*

(c) *There is a sequence in S converging to x.*

B.23 Proposition. *Let S be a subset of M. Then $\partial S = \overline{S} \setminus S^\circ$. An element x of M is in ∂S if and only if every neighborhood of x intersects both S and $M \setminus S$.*

B.24 Definition. A subset S of M is *complete* if each Cauchy sequence in S has a limit in S. If the entire space M is complete, then d is a *complete metric*.

B.25 Proposition. *Every complete subset of M is closed.*

B.26 Proposition. *Every closed subset of a complete subset of M is itself complete.*

B.27 Definition. Let S be a subset of M. Then S is *compact* if, for each collection $\mathfrak{G}$ of open subsets of M whose union includes S, there is a finite subcollection of $\mathfrak{G}$ whose union includes S. That is, the set S is compact if each open covering of S can be thinned to a finite subcovering.

B.28 Definition. Let S be a subset of M. Consider S to be a metric space with the metric inherited from M. Then S has the *finite intersection property* if $\bigcap\{F : F \in \mathfrak{F}\}$ is nonempty whenever $\mathfrak{F}$ is a collection of closed subsets of S such that each finite subcollection of $\mathfrak{F}$ has nonempty intersection.

In the preceding definition, the set S, not M, should really be considered to be the universal set when taking intersections of empty families $\mathfrak{F}$. However, the property being defined is the same either way.

B.29 Proposition. *Let S be a subset of M. Then the following are equivalent.*

(a) *The set S is compact.*
(b) *The set S has the finite intersection property.*
(c) *Each sequence in S has a convergent subsequence whose limit is in S.*

B.30 Proposition. *Every closed subset of a compact subset of M is itself compact.*

B.31 Proposition. *Every compact subset of M is complete, hence closed.*

B.32 Proposition. *Every compact subset of M is bounded.*

B.33 Definition. Suppose that $D \subseteq S \subseteq M$. Then D is *dense* in S if $\overline{D} \supseteq S$.

B.34 Proposition. *Suppose that $D \subseteq S \subseteq M$. Then the following are equivalent.*

(a) *The set D is dense in S.*
(b) *For every x in S and every positive ϵ, there is a y in D such that $d(x, y) < \epsilon$.*
(c) *Every element of S is the limit of a sequence from D.*

B.35 Definition. A subset of M is *separable* if it has a countable dense subset.

B.36 Proposition. *Every subset of a separable subset of M is itself separable.*

B.37 Proposition. *Every compact subset of M is separable.*

B.38 Definition. Let f be a function from M_1 into M_2. Then f is

(a) *continuous at the point x of M_1* if, for each neighborhood V of $f(x)$, there is a neighborhood U of x such that $f(U) \subseteq V$;

(b) *continuous on M_1* or just *continuous* if, for each open subset V of M_2, the set $f^{-1}(V)$ is open.

B.39 Proposition. *Suppose that f is a function from M_1 into M_2 and that x is an element of M_1. Then the following are equivalent.*

(a) *The function f is continuous at x.*

(b) *For every positive ϵ there is a positive δ_ϵ such that $d_2(f(x), f(y)) < \epsilon$ whenever $y \in M_1$ and $d_1(x, y) < \delta_\epsilon$.*

(c) *Whenever a sequence (x_n) in M_1 converges to x, the sequence $(f(x_n))$ converges to $f(x)$.*

B.40 Proposition. *Let f be a function from M_1 into M_2. Then f is continuous on M_1 if and only if f is continuous at every point of M_1.*

B.41 Proposition. *Suppose that f and g are continuous functions from M_1 into M_2. If f and g agree on a dense subset of M_1, then f and g agree on all of M_1.*

B.42 Proposition. *Suppose that S is a subset of M_1 and that f is a continuous function from M_1 into M_2. If S is compact or separable, then $f(S)$ has that same property.*

B.43 Proposition. *Suppose that S is a nonempty compact subset of M and that f is a continuous real-valued function on M. Then f is bounded on S, and furthermore there are points x_1 and x_2 in S such that $f(x_1) = \inf\{ f(x) : x \in S \}$ and $f(x_2) = \sup\{ f(x) : x \in S \}$.*

B.44 Definition. A function f from M_1 into M_2 is *uniformly continuous on M_1* or just *uniformly continuous* if, for every positive ϵ, there is a positive δ_ϵ such that $d_2(f(x), f(y)) < \epsilon$ whenever $x, y \in M_1$ and $d_1(x, y) < \delta_\epsilon$.

B.45 Proposition. *Every uniformly continuous function from M_1 into M_2 is continuous.*

B.46 Proposition. *If M_1 is compact, then every continuous function from M_1 into M_2 is uniformly continuous.*

B.47 Definition. Let (f_n) be a sequence of functions from M_1 into M_2. Then (f_n) is *uniformly Cauchy* if, for every positive ϵ, there is a positive integer N_ϵ such that $d_2\big(f_n(x), f_m(x)\big) < \epsilon$ whenever $n, m \geq N_\epsilon$ and $x \in M_1$. The sequence (f_n) *converges uniformly* to a function f from M_1 into M_2, and f is called the *uniform limit* of (f_n), if, for every positive ϵ, there is a positive integer N_ϵ such that $d_2\big(f_n(x), f(x)\big) < \epsilon$ whenever $n \geq N_\epsilon$.

B.48 Proposition. *Suppose that (f_n) and f are, respectively, a sequence of functions and a function, all mapping M_1 into M_2. If each f_n is continuous and (f_n) converges uniformly to f, then f is continuous.*

B.49 Corollary. *Suppose that (f_n) is a uniformly Cauchy sequence of continuous functions from M_1 into M_2 and that M_2 is complete. Then there is a continuous function f from M_1 into M_2 to which (f_n) converges uniformly.*

B.50 Definition. For each pair $(x_1, \ldots, x_N), (y_1, \ldots, y_N)$ of elements of the Cartesian product $M_1 \times \cdots \times M_N$, let

$$d_p\big((x_1, \ldots, x_N), (y_1, \ldots, y_N)\big) = \left(\sum_{j=1}^{N} \big(d_j(x_j, y_j)\big)^2\right)^{1/2}.$$

Then d_p is the *product metric* on $M_1 \times \cdots \times M_N$ induced by $d_1, \ldots, d_N$, and the corresponding topology is the *product topology* for $M_1 \times \cdots \times M_N$ induced by $d_1, \ldots, d_N$.

B.51 Proposition. *The product metric on $M_1 \times \cdots \times M_N$ induced by $d_1, \ldots, d_N$ is a metric.*

It is a standard convention that whenever reference is made to an unspecified metric or topology for the Cartesian product of a nonempty finite list of metric spaces, for example by saying that a subset of the Cartesian product is open without indicating the relevant metric or topology, it is the product metric or product topology that is being assumed. In that spirit, all references in the rest of this section to $M_1 \times \cdots \times M_N$ as a metric space imply the metric d_p of Definition B.50.

B.52 Proposition. *Suppose that $\big((x_1^{(n)}, \ldots, x_N^{(n)})\big)_{n=1}^{\infty}$ is a sequence in $M_1 \times \cdots \times M_N$ and that $(x_1, \ldots, x_N) \in M_1 \times \cdots \times M_N$. Then*

$$\lim_n (x_1^{(n)}, \ldots, x_N^{(n)}) = (x_1, \ldots, x_N)$$

if and only if $\lim_n x_j^{(n)} = x_j$ when $j = 1, \ldots, N$.

B.53 Proposition. *A subset of $M_1 \times \cdots \times M_N$ is open if and only if it is the union of sets of the form $U_1 \times \cdots \times U_N$, where each U_j is an open subset of the corresponding M_j.*

B.54 Proposition. *Suppose that $M_1, \ldots, M_N$ are all nonempty. Then $M_1 \times \cdots \times M_N$ is complete, or compact, or separable, or bounded, if and only if each M_j has that same property.*

B.55 Definition. Two metrics on the same set are *(topologically) equivalent* if they induce the same topology.

Clearly, a set that is closed with respect to a metric d is also closed with respect to every metric equivalent to d. Such properties that depend not on the particular metric but rather only on the topology induced by the metric are called *topological properties.* It is easy to see that convergence of a sequence to a particular limit, compactness, separability, and continuity are topological properties. Cauchyness, completeness, and boundedness are not. See Exercise 1.42 for an example of a metric d on $\mathbb{R}$ that is equivalent to the usual metric on $\mathbb{R}$ and yet has the property that $\mathbb{R}$ is bounded and incomplete under d; the incompleteness implies the existence of a sequence of reals that is not Cauchy in the usual sense but that is d-Cauchy.

B.56 Definition. Suppose that f is a one-to-one map from M_1 onto M_2 such that both f and f^{-1} are continuous. Then f is a *homeomorphism,* and M_1 and M_2 are *homeomorphic.*

Notice that two metrics on the same set are equivalent if and only if the identity map on the set, treated as a function between the two metric spaces, is a homeomorphism.

B.57 Proposition. *Suppose that f is a homeomorphism from M_1 onto M_2. Then a subset S of M_1 is open, or closed, or compact, or separable, if and only if $f(S)$ has that same property. A sequence (x_n) in M_1 converges to an element x of M_1 if and only if the sequence $\big(f(x_n)\big)$ converges to $f(x)$ in M_2.*

Basically, a homeomorphism is a one-to-one function from one metric space onto another that preserves topological properties in both directions. Homeomorphisms do not have to preserve Cauchyness, completeness, and boundedness. For example, if $\mathbb{R}$ is treated as a metric space M_1 with its usual metric and as another metric space M_2 with the metric of Exercise 1.42, then the identity function on $\mathbb{R}$ can be viewed as a homeomorphism from M_1 onto M_2 that does not preserve Cauchyness, completeness, and boundedness. See the comments following Definition B.55.

One way to force a homeomorphism to preserve properties closely related to the metric as well as topological properties is to require it to preserve distances.

B.58 Definition. Suppose that f is a map from M_1 onto M_2 such that $d_2\big(f(x), f(y)\big) = d_1(x, y)$ whenever $x, y \in M_1$. Then f is an *isometry,* and M_1 and M_2 are *isometric.*

B.59 Proposition. *Suppose that f is an isometry from M_1 onto M_2. Then f is a homeomorphism. A subset S of M_1 is complete or bounded if and only if $f(S)$ has that same property. A sequence (x_n) in M_1 is Cauchy if and only if the sequence $\big(f(x_n)\big)$ in M_2 is Cauchy.*

B.60 Definition. The metric space M_2 is a *completion* of M_1 if there is an isometry from M_1 onto a dense subset of M_2.

The following result can be proved by isometrically embedding the metric space M into the metric space of all bounded real-valued continuous functions on M. See, for example, [129, pp. 90–92].

B.61 Theorem. *Every metric space has a completion.*

Appendix C

The Spaces ℓ_p and ℓ_p^n, $1 \le p \le \infty$

The spaces ℓ_p and ℓ_p^n, where $1 \le p \le \infty$ and $n \in \mathbb{N}$, are treated in Chapter 1 as the Lebesgue spaces $L_p(\Omega, \Sigma, \mu)$, where Ω is $\mathbb{N}$ or the set $\{1, \dots, n\}$ and μ is the counting measure on the σ-algebra Σ of all subsets of Ω. The purpose of this section is to provide a more elementary development of the spaces ℓ_p and ℓ_p^n paralleling that given in Chapter 1 but requiring only as much knowledge of analysis as would come from a first undergraduate course in real analysis without measure theory, and no knowledge of general topology beyond the basic facts about metric spaces from Appendix B. For the sake of completeness, a few words will also be said about the trivial normed space ℓ_p^0.

In the following definition, the p^{th} root of infinity is to be interpreted to be infinity when $1 \le p < \infty$.

C.1 Definition. Suppose that $1 \le p \le \infty$ and that X is the vector space of all sequences of scalars with the usual vector space operations, that is, with addition of sequences and multiplication of sequences by scalars performed term by term. For each member (α_j) of X, let

$$\|(\alpha_j)\|_p = \begin{cases} \left(\sum_{j=1}^{\infty} |\alpha_j|^p\right)^{1/p} & \text{if } 1 \le p < \infty; \\ \sup\{\, |\alpha_j| : j \in \mathbb{N}\,\} & \text{if } p = \infty. \end{cases}$$

Then the *p-norm* on X is the function $\|\cdot\|_p \colon X \to [0, \infty]$. The collection of all sequences (α_j) of scalars such that $\|(\alpha_j)\|_p$ is finite is denoted by ℓ_p (pronounced "little ell p").

The p-norms are not really norms on the vector space X of the preceding definition. For example, if $\alpha_j = j$ for each j, then $\|(\alpha_j)\|_p = \infty$ when $1 \le p \le \infty$. It does turn out that each of the spaces ℓ_p is a subspace of X and has the corresponding p-norm as a norm. To show this, the first order of business is to obtain some classical inequalities.

C.2 Definition. Suppose that $1 \le p \le \infty$. Define q as follows. If $p = 1$, then let $q = \infty$. If $p = \infty$, then let $q = 1$. If $1 < p < \infty$, then let q be such that $p^{-1} + q^{-1} = 1$; that is, let $q = p/(p-1)$. Then q is the *exponent conjugate to p*, and p and q are *conjugate exponents.*

Notice that $1 \le q \le \infty$ for each conjugate exponent q, and that if q is conjugate to p, then p is conjugate to q. Notice also that $q = 2$ when $p = 2$, and that 2 is the only value of p that is its own conjugate exponent.

C.3 Lemma. *Suppose that $1 < p < \infty$ and that q is conjugate to p. Then*

$$rs \le \frac{r^p}{p} + \frac{s^q}{q}$$

for all nonnegative real numbers r and s.

PROOF. Suppose that $0 < \alpha < 1$ and that $f(t) = t^\alpha - \alpha t$ when $t > 0$. It is a straightforward calculus exercise to show that f takes on its maximum value when $t = 1$, so $t^\alpha - \alpha t \le 1 - \alpha$ when $t > 0$. Substituting the quotient u/v of two positive numbers for t, multiplying by v, and rearranging the resulting inequality shows that

$$u^\alpha v^{1-\alpha} \le \alpha u + (1-\alpha)v$$

when $u, v > 0$. It is clear that this inequality holds, in fact, when $u, v \ge 0$. Finally, let r and s be nonnegative reals, let $u = r^p$, let $v = s^q$, let $\alpha = p^{-1}$ (so that $1 - \alpha = q^{-1}$), and substitute into the above inequality to finish the proof. ■

The interpretation of one special case of the following result relies on the usual convention that $0 \cdot \infty = 0$.

C.4 Hölder's Inequality for Sequences. *Suppose that $1 \le p \le \infty$ and that q is conjugate to p. Then*

$$\|(\alpha_j \beta_j)\|_1 \le \|(\alpha_j)\|_p \, \|(\beta_j)\|_q$$

for all sequences (α_j) and (β_j) of scalars.

PROOF. It may be assumed that neither (α_j) nor (β_j) is the zero sequence, and therefore that $\|(\alpha_j)\|_p$ and $\|(\beta_j)\|_q$ are both nonzero. It may also be assumed that $\|(\alpha_j)\|_p$ and $\|(\beta_j)\|_q$ are finite, and therefore are positive reals. Now suppose that the conclusion of Hölder's inequality holds when $\|(\alpha_j)\|_p = \|(\beta_j)\|_q = 1$. For the general case, let $\gamma = \|(\alpha_j)\|_p$ and $\delta = \|(\beta_j)\|_q$, so that $\gamma, \delta \in (0, \infty)$. It is easy to check that

$$\|(\gamma^{-1}\alpha_j)\|_p = \|(\delta^{-1}\beta_j)\|_q = 1,$$

which implies that

$$\begin{aligned}\|(\alpha_j)\|_p^{-1}\|(\beta_j)\|_q^{-1}\|(\alpha_j\beta_j)\|_1 &= \|(\gamma^{-1}\alpha_j\delta^{-1}\beta_j)\|_1 \\ &\le \|(\gamma^{-1}\alpha_j)\|_p\,\|(\delta^{-1}\beta_j)\|_q \\ &= 1.\end{aligned}$$

Multiplying through by $\|(\alpha_j)\|_p\,\|(\beta_j)\|_q$ then yields the desired result. It may therefore be assumed that $\|(\alpha_j)\|_p = \|(\beta_j)\|_q = 1$. The proof will be complete once it is shown that $\|(\alpha_j\beta_j)\|_1 \le 1$.

Suppose first that either p or q is 1. It may be assumed that $p = 1$ and therefore that $q = \infty$. Since $\|(\beta_j)\|_\infty = 1$, it follows that $|\beta_j| \le 1$ for each j, and therefore that

$$\|(\alpha_j\beta_j)\|_1 = \sum_j |\alpha_j\beta_j| \le \sum_j |\alpha_j| = \|(\alpha_j)\|_1 = 1,$$

as required.

Finally, suppose that $1 < p < \infty$. It follows from the lemma that

$$\begin{aligned}\|(\alpha_j\beta_j)\|_1 &= \sum_j |\alpha_j\beta_j| \\ &\le \sum_j \left(\frac{|\alpha_j|^p}{p} + \frac{|\beta_j|^q}{q}\right) \\ &= \frac{\|(\alpha_j)\|_p^p}{p} + \frac{\|(\beta_j)\|_q^q}{q} \\ &= \frac{1}{p} + \frac{1}{q} \\ &= 1,\end{aligned}$$

which finishes the proof. ■

In a sense, Hölder's inequality is just a generalization of Cauchy's inequality for the spaces $\mathbb{F}^n$ such that $n \in \mathbb{N}$, and in fact Cauchy's inequality follows easily from Hölder's. Suppose that $n \in \mathbb{N}$, that $(\alpha_1, \dots, \alpha_n)$ and $(\beta_1, \dots, \beta_n)$ are ordered n-tuples of scalars, and that (α_j) and (β_j) are the corresponding extensions of $(\alpha_1, \dots, \alpha_n)$ and $(\beta_1, \dots, \beta_n)$ to sequences of scalars formed by letting $\alpha_j = \beta_j = 0$ when $j > n$. Then Hölder's inequality with p set to 2 implies that

$$\begin{aligned}\left|\sum_{j=1}^{n} \alpha_j\beta_j\right| &\le \sum_{j=1}^{n} |\alpha_j\beta_j| \\ &= \|(\alpha_j\beta_j)\|_1 \\ &\le \|(\alpha_j)\|_2 \, \|(\beta_j)\|_2 \\ &= \left(\sum_{j=1}^{n} |\alpha_j|^2\right)^{1/2} \left(\sum_{j=1}^{n} |\beta_j|^2\right)^{1/2},\end{aligned}$$

which is the conclusion of Cauchy's inequality.

Just as the triangle inequality for the norm of Euclidean n-space follows from Cauchy's inequality, a triangle inequality for the p-norm follows from Hölder's inequality.

C.5 Minkowski's Inequality for Sequences. *Suppose that* $1 \le p \le \infty$. *Then*

$$\|(\alpha_j) + (\beta_j)\|_p \le \|(\alpha_j)\|_p + \|(\beta_j)\|_p$$

for all sequences (α_j) *and* (β_j) *of scalars.*

PROOF. If $p = 1$, then

$$\|(\alpha_j) + (\beta_j)\|_1 = \sum_j |\alpha_j + \beta_j| \le \sum_j \left(|\alpha_j| + |\beta_j|\right) = \|(\alpha_j)\|_1 + \|(\beta_j)\|_1.$$

If $p = \infty$, then $|\alpha_k + \beta_k| \le |\alpha_k| + |\beta_k| \le \|(\alpha_j)\|_\infty + \|(\beta_j)\|_\infty$ for each k, so $\|(\alpha_j) + (\beta_j)\|_\infty \le \|(\alpha_j)\|_\infty + \|(\beta_j)\|_\infty$.

Finally, suppose that $1 < p < \infty$, and let q be conjugate to p. It may be assumed that $\|(\alpha_j) + (\beta_j)\|_p$ is nonzero and that both $\|(\alpha_j)\|_p$ and $\|(\beta_j)\|_p$ are finite. Since

$$\begin{aligned}\sum_j |\alpha_j + \beta_j|^p &\le \sum_j \left(|\alpha_j| + |\beta_j|\right)^p \\ &\le \sum_j \left(2\max\{|\alpha_j|, |\beta_j|\}\right)^p \\ &= 2^p \sum_j \max\{|\alpha_j|^p, |\beta_j|^p\} \\ &\le 2^p \left(\sum_j |\alpha_j|^p + \sum_j |\beta_j|^p\right),\end{aligned}$$

the finiteness of $\|(\alpha_j)\|_p$ and $\|(\beta_j)\|_p$ implies that of $\|(\alpha_j)+(\beta_j)\|_p$, and so $0 < \|(\alpha_j)+(\beta_j)\|_p < \infty$. Next, notice that

$$\begin{aligned}\|(\alpha_j)+(\beta_j)\|_p^p &= \sum_j |\alpha_j+\beta_j|^p \\ &= \sum_j |\alpha_j+\beta_j|\,|\alpha_j+\beta_j|^{p-1} \\ &\le \sum_j |\alpha_j|\,|\alpha_j+\beta_j|^{p-1} + \sum_j |\beta_j|\,|\alpha_j+\beta_j|^{p-1}.\end{aligned}$$

Now apply Hölder's inequality to the sequences $(|\alpha_j|)$ and $(|\alpha_j+\beta_j|^{p-1})$, which yields the inequality

$$\begin{aligned}\sum_j |\alpha_j|\,|\alpha_j+\beta_j|^{p-1} &\le \Big(\sum_j |\alpha_j|^p\Big)^{1/p}\Big(\sum_j |\alpha_j+\beta_j|^{(p-1)q}\Big)^{1/q} \\ &= \Big(\sum_j |\alpha_j|^p\Big)^{1/p}\Big(\sum_j |\alpha_j+\beta_j|^{p}\Big)^{(p-1)/p} \\ &= \|(\alpha_j)\|_p\,\|(\alpha_j)+(\beta_j)\|_p^{p-1}.\end{aligned}$$

Since an analogous inequality holds for $\sum_j |\beta_j|\,|\alpha_j+\beta_j|^{p-1}$,

$$\|(\alpha_j)+(\beta_j)\|_p^p \le \big(\|(\alpha_j)\|_p + \|(\beta_j)\|_p\big)\|(\alpha_j)+(\beta_j)\|_p^{p-1}.$$

Dividing both sides by $\|(\alpha_j)+(\beta_j)\|_p^{p-1}$ yields the desired result. ■

C.6 Theorem. *Suppose that $1 \le p \le \infty$. Then ℓ_p is a vector space when the sum of two sequences of scalars and the product of a scalar and a sequence of scalars are defined in the usual way. The p-norm is a norm on this vector space.*

PROOF. It is easy to check that $\alpha(\alpha_j) \in \ell_p$ whenever $(\alpha_j) \in \ell_p$ and α is a scalar. Also, each sum of two elements of ℓ_p is in ℓ_p by Minkowski's inequality. Since the zero sequence is in ℓ_p, it follows that ℓ_p is a subspace of the vector space of all sequences of scalars with the usual operations. Minkowski's inequality provides a triangle inequality for the p-norm, while the other properties that the p-norm must have to be a norm on ℓ_p follow quickly from the definitions. ■

Henceforth, when $1 \le p \le \infty$ and ℓ_p is treated as a normed space without the norm being specified, the p-norm is implied.

C.7 Definition. Suppose that $1 \le p \le \infty$ and that n is a positive integer. For each member $(\alpha_1,\dots,\alpha_n)$ of $\mathbb{F}^n$, let

$$\|(\alpha_1,\dots,\alpha_n)\|_p = \begin{cases}\Big(\sum_{j=1}^n |\alpha_j|^p\Big)^{1/p} & \text{if } 1 \le p < \infty;\\ \max\{|\alpha_1|,\dots,|\alpha_n|\} & \text{if } p = \infty.\end{cases}$$

Then the *p-norm* on $\mathbb{F}^n$ is the function $\|\cdot\|_p \colon \mathbb{F}^n \to [0,\infty)$. By convention, the p-norm is defined on the zero-dimensional vector space $\mathbb{F}^0$ by the formula $\|0\|_p = 0$.

C.8 Theorem. *Suppose that $1 \le p \le \infty$ and that n is a nonnegative integer. Then the p-norm is a norm on the vector space $\mathbb{F}^n$.*

PROOF. This is obvious when $n = 0$, so it will be assumed that $n \ge 1$. Extend each element $(\alpha_1, \ldots, \alpha_n)$ of $\mathbb{F}^n$ to an element (α_j) of ℓ_p by letting $\alpha_j = 0$ when $j > n$. Then

$$\begin{aligned} \|(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n)\|_p &= \|(\alpha_j) + (\beta_j)\|_p \\ &\le \|(\alpha_j)\|_p + \|(\beta_j)\|_p \\ &= \|(\alpha_1, \ldots, \alpha_n)\|_p + \|(\beta_1, \ldots, \beta_n)\|_p \end{aligned}$$

whenever $(\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n) \in \mathbb{F}^n$, so the p-norm satisfies the triangle inequality. It is easily checked that the p-norm has the other properties required of a norm. ∎

C.9 Definition. Suppose that $1 \le p \le \infty$ and that n is a nonnegative integer. Then ℓ_p^n (pronounced "little ell p n") is the normed space formed by the vector space $\mathbb{F}^n$ with the p-norm.

C.10 Theorem. *Suppose that $1 \le p \le \infty$. Then ℓ_p is a Banach space, as is ℓ_p^n for each nonnegative integer n.*

PROOF. Let $((\alpha_j^{(k)}))_{k=1}^\infty$ be a Cauchy sequence of elements of ℓ_p. Since

$$|\alpha_{j_0}^{(k)} - \alpha_{j_0}^{(l)}| \le \|(\alpha_j^{(k)}) - (\alpha_j^{(l)})\|_p$$

whenever $j_0, k, l \in \mathbb{N}$, each of the sequences $(\alpha_{j_0}^{(k)})_{k=1}^\infty$ such that $j_0 \in \mathbb{N}$ is Cauchy, hence convergent in $\mathbb{F}$. For each positive integer j_0, let $\alpha_{j_0} = \lim_k \alpha_{j_0}^{(k)}$. The proof for ℓ_p will be finished once it is shown that $(\alpha_j) \in \ell_p$ and that $\|(\alpha_j^{(k)}) - (\alpha_j)\|_p \to 0$ as $k \to \infty$.

Suppose first that $p = \infty$. Let ϵ be a positive real number and let N_ϵ be a positive integer such that $\|(\alpha_j^{(k)}) - (\alpha_j^{(l)})\|_\infty \le \epsilon$ when $k, l \ge N_\epsilon$. It follows that $|\alpha_{j_0}^{(k)} - \alpha_{j_0}^{(l)}| \le \epsilon$ for each j_0 when $k, l \ge N_\epsilon$, and therefore that

$$|\alpha_{j_0}^{(k)} - \alpha_{j_0}| = \lim_l |\alpha_{j_0}^{(k)} - \alpha_{j_0}^{(l)}| \le \epsilon$$

for each j_0 when $k \ge N_\epsilon$. Taking the supremum over all j_0 shows that $\|(\alpha_j^{(k)}) - (\alpha_j)\|_\infty \le \epsilon$ when $k \ge N_\epsilon$. Then

$$(\alpha_j) = -((\alpha_j^{(N_\epsilon)}) - (\alpha_j)) + (\alpha_j^{(N_\epsilon)}) \in \ell_\infty,$$

and $\|(\alpha_j^{(k)}) - (\alpha_j)\|_\infty \to 0$ as $k \to \infty$.

Now suppose that $1 \le p < \infty$. Let ϵ be a positive real number and let N_ϵ be a positive integer such that $\|(\alpha_j^{(k)}) - (\alpha_j^{(l)})\|_p \le \epsilon$ when $k, l \ge N_\epsilon$. It follows that for each positive integer j_0,

$$\left(\sum_{j=1}^{j_0} |\alpha_j^{(k)} - \alpha_j^{(l)}|^p\right)^{1/p} \le \|(\alpha_j^{(k)}) - (\alpha_j^{(l)})\|_p \le \epsilon$$

when $k, l \ge N_\epsilon$. Leaving j_0 fixed and letting l tend to infinity shows that

$$\left(\sum_{j=1}^{j_0} |\alpha_j^{(k)} - \alpha_j|^p\right)^{1/p} \le \epsilon$$

for each positive integer j_0 when $k \ge N_\epsilon$. Letting j_0 tend to infinity shows that

$$\|(\alpha_j^{(k)}) - (\alpha_j)\|_p = \left(\sum_{j=1}^{\infty} |\alpha_j^{(k)} - \alpha_j|^p\right)^{1/p} \le \epsilon$$

when $k \ge N_\epsilon$. Therefore $(\alpha_j) \in \ell_p$ by the argument used for ℓ_∞, and $\|(\alpha_j^{(k)}) - (\alpha_j)\|_p \to 0$ as $k \to \infty$. This completes the proof for ℓ_p.

Suppose that n is a positive integer and that $\left((\alpha_1^{(k)}, \ldots, \alpha_n^{(k)})\right)_{k=1}^{\infty}$ is a Cauchy sequence in ℓ_p^n. Then each sequence $(\alpha_j^{(k)})_{k=1}^{\infty}$ such that $j = 1, \ldots, n$ is a Cauchy sequence of scalars and so has a limit α_j. Depending on whether or not p is finite, either

$$\|(\alpha_1^{(k)}, \ldots, \alpha_n^{(k)}) - (\alpha_1, \ldots, \alpha_n)\|_p = \left(\sum_{j=1}^{n} |\alpha_j^{(k)} - \alpha_j|^p\right)^{1/p} \to 0$$

or

$$\|(\alpha_1^{(k)}, \ldots, \alpha_n^{(k)}) - (\alpha_1, \ldots, \alpha_n)\|_\infty = \max\left\{|\alpha_1^{(k)} - \alpha_1|, \ldots, |\alpha_n^{(k)} - \alpha_n|\right\} \to 0$$

as $k \to \infty$, and so

$$\lim_k (\alpha_1^{(k)}, \ldots, \alpha_n^{(k)}) = (\alpha_1, \ldots, \alpha_n).$$

This proves that ℓ_p^n is complete. Since ℓ_p^0 is obviously complete, the theorem is proved. ■

C.11 Corollary. *For each nonnegative integer n, Euclidean n-space is a Banach space.*

The following two results are necessary for Examples 1.10.3 and 1.10.5 if the duality theory for ℓ_p and ℓ_p^n is not to be obtained from the more general duality theory for Lebesgue spaces.

C.12 Theorem. *Suppose that $1 \le p < \infty$ and that q is conjugate to p. For each element (β_j) of ℓ_q, let $T(\beta_j)$ be the scalar-valued function on ℓ_p defined by the formula*

$$\big(T(\beta_j)\big)(\alpha_j) = \sum_j \alpha_j \beta_j.$$

Then T is an isometric isomorphism from ℓ_q onto ℓ_p^*.

PROOF. If $(\alpha_j) \in \ell_p$ and $(\beta_j) \in \ell_q$, then

$$\sum_j |\alpha_j \beta_j| \le \|(\alpha_j)\|_p \, \|(\beta_j)\|_q < \infty$$

by Hölder's inequality, so $\sum_j \alpha_j \beta_j$ converges. This shows that the definition of T makes sense and that

$$\big|\big(T(\beta_j)\big)(\alpha_j)\big| \le \|(\alpha_j)\|_p \, \|(\beta_j)\|_q \tag{C.1}$$

whenever $(\alpha_j) \in \ell_p$ and $(\beta_j) \in \ell_q$. For each (β_j) in ℓ_q, the function $T(\beta_j)$ is clearly a linear functional on ℓ_p, and (C.1) implies that this linear functional is bounded and satisfies the inequality

$$\|T(\beta_j)\| \le \|(\beta_j)\|_q. \tag{C.2}$$

It is easy to check that the function $T\colon \ell_q \to \ell_p^*$ is a linear operator. The theorem will be proved once it is shown that T maps ℓ_q onto ℓ_p^* and that equality holds in (C.2) whenever $(\beta_j) \in \ell_q$.

Suppose that $x^* \in \ell_p^*$. For each standard unit vector e_j of ℓ_p, let $\beta_j = x^* e_j$ and let γ_j be a scalar such that $|\gamma_j| = 1$ and $|\beta_j| = \gamma_j \beta_j$. Suppose for the moment that $p \ne 1$, so that $1 < q < \infty$. Then for each positive integer k,

$$\begin{aligned}
\sum_{j=1}^k |\beta_j|^q &= \sum_{j=1}^k \gamma_j \beta_j |\beta_j|^{q-1} \\
&= x^*\Bigg(\sum_{j=1}^k \gamma_j |\beta_j|^{q-1} e_j\Bigg) \\
&\le \|x^*\| \, \Bigg\|\sum_{j=1}^k \gamma_j |\beta_j|^{q-1} e_j\Bigg\|_p \\
&= \|x^*\| \Bigg(\sum_{j=1}^k |\beta_j|^{(q-1)p}\Bigg)^{1/p} \\
&= \|x^*\| \Bigg(\sum_{j=1}^k |\beta_j|^q\Bigg)^{1/p},
\end{aligned}$$

which implies that

$$\left(\sum_{j=1}^{k} |\beta_j|^q\right)^{1/q} = \left(\sum_{j=1}^{k} |\beta_j|^q\right)^{1-1/p} \le \|x^*\|.$$

Letting k tend to infinity shows that $(\beta_j) \in \ell_q$ and that

$$\|(\beta_j)\|_q \le \|x^*\|. \tag{C.3}$$

Now suppose that $p = 1$, so that $q = \infty$. For each positive integer j,

$$|\beta_j| = |x^* e_j| \le \|x^*\| \, \|e_j\|_1 = \|x^*\|,$$

which, by taking the supremum over all j, again shows that $(\beta_j) \in \ell_q$ and that (C.3) holds. Thus, whether or not $p = 1$, it follows that $(\beta_j) \in \ell_q$ and therefore that $T(\beta_j) \in \ell_p^*$. For every element (α_j) of ℓ_p,

$$x^*(\alpha_j) = x^*\left(\sum_j \alpha_j e_j\right) = \sum_j \alpha_j x^* e_j = \sum_j \alpha_j \beta_j = \big(T(\beta_j)\big)(\alpha_j),$$

so $T(\beta_j) = x^*$. The operator T therefore maps ℓ_q onto ℓ_p^*.

All that remains to be shown is that $\|T(\beta_j)\| = \|(\beta_j)\|_q$ when $(\beta_j) \in \ell_q$. First notice that $T(\beta_j)$ is nonzero if (β_j) is a nonzero element of ℓ_q, so T is one-to-one. Now fix an element (β_j) of ℓ_q and let $x^* = T(\beta_j)$. The argument of the preceding paragraph yields an element (β_j') of ℓ_q such that $T(\beta_j') = x^*$, and (β_j') must equal (β_j) since T is one-to-one. It follows from (C.3) that

$$\|(\beta_j)\|_q \le \|T(\beta_j)\|. \tag{C.4}$$

Combining (C.2) and (C.4) yields the desired equality. ■

C.13 Theorem. *Suppose that $1 \le p \le \infty$, that q is conjugate to p, and that n is a positive integer. For each element $(\beta_1, \dots, \beta_n)$ of ℓ_q^n, let $T(\beta_1, \dots, \beta_n)$ be the scalar-valued function on ℓ_p^n defined by the formula*

$$\big(T(\beta_1, \dots, \beta_n)\big)(\alpha_1, \dots, \alpha_n) = \sum_{j=1}^{n} \alpha_j \beta_j.$$

Then T is an isometric isomorphism from ℓ_q^n onto $(\ell_p^n)^*$.

PROOF. For each element $(\beta_1, \dots, \beta_n)$ of ℓ_q^n, the map $T(\beta_1, \dots, \beta_n)$ is clearly a linear functional on the finite-dimensional normed space ℓ_p^n, and so is automatically bounded. It follows that T is a linear operator from ℓ_q^n into $(\ell_p^n)^*$. It is easy to check that $\ker(T) = \{(0, \dots, 0)\}$, so T is one-to-one. Since ℓ_q^n and $(\ell_p^n)^*$ both have dimension n, the operator T maps ℓ_q^n onto $(\ell_p^n)^*$.

Let $(\beta_1, \dots, \beta_n)$ be an element of ℓ_q^n. All that remains to be shown is that $\|T(\beta_1, \dots, \beta_n)\| = \|(\beta_1, \dots, \beta_n)\|_q$. Suppose that $(\alpha_1, \dots, \alpha_n) \in \ell_p^n$ and that $(\alpha_1, \dots, \alpha_n)$ and $(\beta_1, \dots, \beta_n)$ are extended to the respective sequences (α_j) and (β_j) by letting $\alpha_j = \beta_j = 0$ when $j > n$. Hölder's inequality implies that

$$\begin{aligned} \big|(T(\beta_1, \dots, \beta_n))(\alpha_1, \dots, \alpha_n)\big| &\le \sum_{j=1}^{n} |\alpha_j \beta_j| \\ &\le \|(\alpha_j)\|_p \, \|(\beta_j)\|_q \\ &= \|(\alpha_1, \dots, \alpha_n)\|_p \, \|(\beta_1, \dots, \beta_n)\|_q, \end{aligned}$$

and so

$$\|T(\beta_1, \dots, \beta_n)\| \le \|(\beta_1, \dots, \beta_n)\|_q. \tag{C.5}$$

Now let $\{e_1, \dots, e_n\}$ be the standard basis for the vector space $\mathbb{F}^n$ and let $\gamma_1, \dots, \gamma_n$ be scalars of absolute value 1 such that $|\beta_j| = \gamma_j \beta_j$ for each j. If $1 < p < \infty$, then

$$\begin{aligned} \sum_{j=1}^{n} |\beta_j|^q &= \sum_{j=1}^{n} \gamma_j \beta_j |\beta_j|^{q-1} \\ &= \big(T(\beta_1, \dots, \beta_n)\big)\big(\gamma_1 |\beta_1|^{q-1}, \dots, \gamma_n |\beta_n|^{q-1}\big) \\ &\le \|T(\beta_1, \dots, \beta_n)\| \, \|(\gamma_1 |\beta_1|^{q-1}, \dots, \gamma_n |\beta_n|^{q-1})\|_p \\ &= \|T(\beta_1, \dots, \beta_n)\| \Big(\sum_{j=1}^{n} |\beta_j|^{(q-1)p} \Big)^{1/p} \\ &= \|T(\beta_1, \dots, \beta_n)\| \Big(\sum_{j=1}^{n} |\beta_j|^{q} \Big)^{1/p}, \end{aligned}$$

which implies that

$$\|(\beta_1, \dots, \beta_n)\|_q = \Big(\sum_{j=1}^{n} |\beta_j|^q \Big)^{1/q} = \Big(\sum_{j=1}^{n} |\beta_j|^q \Big)^{1-1/p} \le \|T(\beta_1, \dots, \beta_n)\|.$$

If $p = 1$, then

$$\begin{aligned} \|(\beta_1, \dots, \beta_n)\|_\infty &= \max\{|\beta_1|, \dots, |\beta_n|\} \\ &= \max\big\{\big|(T(\beta_1, \dots, \beta_n))(e_1)\big|, \dots, \big|(T(\beta_1, \dots, \beta_n))(e_n)\big|\big\} \\ &\le \|T(\beta_1, \dots, \beta_n)\|. \end{aligned}$$

If $p = \infty$, then

$$\begin{aligned}\|(\beta_1, \dots, \beta_n)\|_1 &= \sum_{j=1}^{n} \gamma_j \beta_j \\ &= \big(T(\beta_1, \dots, \beta_n)\big)(\gamma_1, \dots, \gamma_n) \\ &\le \|T(\beta_1, \dots, \beta_n)\| \, \|(\gamma_1, \dots, \gamma_n)\|_\infty \\ &= \|T(\beta_1, \dots, \beta_n)\|.\end{aligned}$$

Therefore $\|(\beta_1, \dots, \beta_n)\|_q \le \|T(\beta_1, \dots, \beta_n)\|$ whatever the value of p, which when combined with (C.5) yields the equality needed to finish the proof. ■

This next result should replace Theorem 1.11.10 if the reflexivity of ℓ_p when $1 < p < \infty$ is not to be obtained from the theory of Lebesgue spaces.

C.14 Theorem. *Suppose that $1 < p < \infty$. Then ℓ_p is reflexive.*

Proof. Let q be conjugate to p, let $T_q \colon \ell_q \to \ell_p^*$ and $T_p \colon \ell_p \to \ell_q^*$ be the usual isometric isomorphisms as in Theorem C.12, and let $Q \colon \ell_p \to \ell_p^{**}$ be the natural map. Suppose that $x^{**} \in \ell_p^{**}$. Then $x^{**}T_q \in \ell_q^*$, so there is a member (α_j) of ℓ_p such that $x^{**}T_q = T_p(\alpha_j)$. If $x^* \in \ell_p^*$, then there is a member (β_j) of ℓ_q such that $x^* = T_q(\beta_j)$, so

$$x^{**}x^* = x^{**}T_q(\beta_j) = \big(T_p(\alpha_j)\big)(\beta_j) = \sum_j \alpha_j \beta_j = \big(T_q(\beta_j)\big)(\alpha_j) = x^*(\alpha_j).$$

It follows that $x^{**} = Q(\alpha_j)$, so Q maps ℓ_p onto ℓ_p^{**}. ■

Appendix D
Ultranets

This appendix is an optional addendum to Section 2.1. Except for Proposition D.10, which requires Proposition 2.1.40 and the material on topological groups preceding it, this appendix uses no material past Proposition 2.1.37.

In addition to the characterizations of compactness and relative compactness in terms of the behavior of nets given in Propositions 2.1.37 and 2.1.40, there are further characterizations of these properties in terms of the behavior of special nets called *ultranets* that can be particularly useful in simplifying compactness arguments involving the axiom of choice. For the definition of ultranets, it is useful to have the following two terms.

D.1 Definition. A net $(x_\alpha)_{\alpha \in I}$ is said to *frequent* a set S or to be in S *frequently* if, for each α in I, there is a β_α in I such that $\alpha \preceq \beta_\alpha$ and $x_{\beta_\alpha} \in S$. The net is said to be in the set *ultimately* if there is an α in I such that $x_\beta \in S$ whenever $\alpha \preceq \beta$.

These terms could have been used to define net convergence and accumulation, since a net converges to (respectively, accumulates at) a point if and only if the net is ultimately (respectively, frequently) in U whenever U is a neighborhood of the point. However, these terms would not have contributed much to the discussion of convergence and accumulation, while they will greatly streamline the presentation of ultranets that is to follow.

Notice that a net is in a set S frequently if and only if it is not ultimately in the complement of S. Clearly, every net has the property that if it is ultimately in a set, then it must be frequently in that set. Ultranets are nets having the converse property.

D.2 Definition. An *ultranet* or *universal net* or *maximal net* is a net with the property that if it is frequently in a set, then it must be ultimately in that set.

Suppose that (x_α) is a net in a set X. If $S \subseteq X$, then (x_α) must frequent either S or $X \setminus S$, from which it follows immediately that an ultranet in X must be either ultimately in S or ultimately in $X \setminus S$. Conversely, suppose that (x_α) has the property that for each subset S of X, the net is either ultimately in S or ultimately in $X \setminus S$. Since no net can be frequently in a set and ultimately in the complement of that set, it follows that (x_α) is ultimately in a set whenever it is frequently in that set, and so is an ultranet. This gives the following result, which is often used to define ultranets.

D.3 Proposition. *A net in a set X is an ultranet if and only if it has the following property: For each subset S of X, the net is either ultimately in S or ultimately in $X \setminus S$.*

D.4 Example. Let x be an element of a set X and let $\mathfrak{M}_x$ be the collection of all subsets of X containing x. Let I be the set of all ordered pairs (a, A) such that $A \in \mathfrak{M}_x$ and $a \in A$, directed by declaring that $(a, A) \preceq (b, B)$ whenever $A \supseteq B$. Let $x_{(a,A)} = a$ whenever $(a, A) \in I$. Then $(x_{(a,A)})$ is a net in X. Notice that for each subset S of X, the net $(x_{(a,A)})$ lies either entirely inside or entirely outside S from some term onward, depending on whether or not $x \in S$. The net $(x_{(a,A)})$ is therefore an ultranet.

It will be useful to study ultranets a bit before pursuing their relationship to compactness. The following result is an immediate consequence of the fact that an ultranet in a topological space that is in a neighborhood of a point frequently must be in that neighborhood ultimately.

D.5 Proposition. *If an ultranet in a topological space accumulates at a point, then it converges to that point.*

Functions preserve ultranets.

D.6 Proposition. *Suppose that (x_α) is an ultranet in a set X and that f is a function from X into a set Y. Then $\big(f(x_\alpha)\big)$ is an ultranet in Y.*

PROOF. Suppose that the net $\big(f(x_\alpha)\big)$ is frequently in the subset S of Y. Then (x_α) is frequently in $f^{-1}(S)$, hence ultimately in $f^{-1}(S)$, and so $\big(f(x_\alpha)\big)$ is ultimately in S. ■

It turns out that every net has a subnet that is an ultranet. The proof of that to be given here depends on the following technical lemma. Readers familiar with the theory of filters will recognize this lemma as a statement about ultrafilters.

D.7 Lemma. *Suppose that (x_α) is a net in a set X. Then there is a family $\mathfrak{M}$ of subsets of X such that*

(1) *the net (x_α) frequents every member of $\mathfrak{M}$;*

(2) *if $M_1, M_2 \in \mathfrak{M}$, then $M_1 \cap M_2 \in \mathfrak{M}$;*

(3) *if $S \subseteq X$, then either S or $X \setminus S$ is in $\mathfrak{M}$.*

PROOF. Let $\mathfrak{A}$ be the collection of all families $\mathfrak{F}$ of subsets of X that satisfy (1) and (2) when "$\mathfrak{M}$" is replaced by "$\mathfrak{F}$." Then $\mathfrak{A}$ is nonempty since $\{X\} \in \mathfrak{A}$. Let $\preceq_{\mathfrak{A}}$ be the preorder on $\mathfrak{A}$ obtained by declaring that $\mathfrak{F}_1 \preceq_{\mathfrak{A}} \mathfrak{F}_2$ when $\mathfrak{F}_1 \subseteq \mathfrak{F}_2$. Then every nonempty chain in $\mathfrak{A}$ has an upper bound, namely, the union of the members of the chain, and of course $\{X\}$ is an upper bound for the empty chain. By Zorn's lemma, there is a maximal element $\mathfrak{M}$ in $\mathfrak{A}$. Notice that $\mathfrak{M} \neq \varnothing$, since $\varnothing \preceq_{\mathfrak{A}} \{X\}$ but $\{X\} \not\preceq_{\mathfrak{A}} \varnothing$. To finish the proof, it is enough to show that $\mathfrak{M}$ satisfies (3).

Let $\mathfrak{M}_0$ be the collection of all subsets of X that have a member of $\mathfrak{M}$ included in them. It is easy to see that $\mathfrak{M}_0 \in \mathfrak{A}$ and that $\mathfrak{M} \preceq_{\mathfrak{A}} \mathfrak{M}_0$, from which it follows that $\mathfrak{M} = \mathfrak{M}_0$. That is, the family $\mathfrak{M}$ contains with each of its elements every superset of that element.

For the rest of this proof, let S be a fixed subset of X. The proof will be done once it is shown that either S or $X \setminus S$ is in $\mathfrak{M}$. To this end, suppose that $M_1, M_2 \in \mathfrak{M}$ and that (x_α) frequents neither $M_1 \cap S$ nor $M_2 \setminus S$. Then (x_α) frequents neither $(M_1 \cap M_2) \cap S$ nor $(M_1 \cap M_2) \setminus S$, so (x_α) does not frequent $M_1 \cap M_2$. Since $M_1 \cap M_2 \in \mathfrak{M}$, this is a contradiction.

Thus, the net (x_α) either frequents every member of $\{\, M \cap S : M \in \mathfrak{M} \,\}$ or frequents every member of $\{\, M \setminus S : M \in \mathfrak{M} \,\}$. By exchanging the roles of S and its complement if necessary, it can be assumed that (x_α) frequents every member of $\{\, M \cap S : M \in \mathfrak{M} \,\}$. Let $\mathfrak{M}_1 = \mathfrak{M} \cup \{\, M \cap S : M \in \mathfrak{M} \,\}$. It is easy to check that $\mathfrak{M}_1 \in \mathfrak{A}$. Since $\mathfrak{M} \preceq_{\mathfrak{A}} \mathfrak{M}_1$, it follows that $\mathfrak{M} = \mathfrak{M}_1$. Finally, let M be any member of the nonempty set $\mathfrak{M}$ and observe that $\mathfrak{M}$ must contain every superset of its element $M \cap S$. It follows that $S \in \mathfrak{M}$. ■

D.8 Theorem. *Every net has a subnet that is an ultranet.*

PROOF. Let $(x_\alpha)_{\alpha \in I}$ be a net in a set X and let $\mathfrak{M}$ be a family of subsets of X such that $\mathfrak{M}$ has the properties listed in the preceding lemma. Let $J = \{\, (\alpha, M) : M \in \mathfrak{M}, \alpha \in I, x_\alpha \in M \,\}$ with the relation $\preceq_J$ given by declaring that $(\alpha_1, M_1) \preceq_J (\alpha_2, M_2)$ when $\alpha_1 \preceq \alpha_2$ and $M_1 \supseteq M_2$. It is easy to check that J with the relation $\preceq_J$ is a directed set and that $\{\, \alpha : (\alpha, M) \in J \,\}$ is cofinal in I. Let $g(\alpha, M) = \alpha$ for each (α, M) in J. Then $(x_{g(\alpha,M)})$ is a subnet of (x_α). If $M_0 \in \mathfrak{M}$, then the fact that (x_α) frequents M_0 produces an α_0 such that $(\alpha_0, M_0) \in J$, and the definition of $\preceq_J$ then guarantees that $x_{g(\alpha,M)} \in M_0$ whenever $(\alpha_0, M_0) \preceq_J (\alpha, M)$. Thus, for each member M_0 of $\mathfrak{M}$, the net $(x_{g(\alpha,M)})$ is ultimately in M_0. Since each subset of X either is in $\mathfrak{M}$ or has its complement in $\mathfrak{M}$, it follows from Proposition D.3 that $(x_{g(\alpha,M)})$ is an ultranet. ■

The next two results are the promised characterizations of compactness and relative compactness in terms of ultranet behavior. The proof of the second is essentially the same as that of the first with the obvious modifications and with the references to Proposition 2.1.37 replaced by references to Proposition 2.1.40. Both results use the axiom of choice in their proofs, since the proofs are based on Theorem D.8 which in turn is obtained from a lemma that uses Zorn's lemma in its proof.

D.9 Proposition. *A subset S of a topological space is compact if and only if each ultranet in S has a limit in S.*

PROOF. If S is compact, then Proposition 2.1.37 assures that every ultranet in S has an accumulation point in S and so has a limit in S. Conversely, suppose that every ultranet in S has a limit in S. Since every net in S has a subnet that is an ultranet, every net in S has a subnet with a limit in S, and so another application of Proposition 2.1.37 shows that S is compact. ■

Notice that the following result is stated only for topological groups.

D.10 Proposition. *A subset S of a topological group is relatively compact if and only if each ultranet in S has a limit in X.*

As an application of the results of this section, a very short proof of Alexander's subbasis theorem will now be given.

D.11 Alexander's Subbasis Theorem. (J. W. Alexander, 1939 [4]). *Suppose that $\mathfrak{S}$ is a subbasis for the topology of a topological space X and that S is a subset of X. If every covering of S by elements of $\mathfrak{S}$ can be thinned to a finite subcovering, then S is compact.*

PROOF. Suppose that S is not compact. Then some ultranet (x_α) in S has no limit in S. It follows from Proposition 2.1.15 that for each x in S there is a member U_x of $\mathfrak{S}$ containing x such that (x_α) is ultimately in $X \setminus U_x$. Let $\mathfrak{C} = \{ U_x : x \in S \}$. Then $\mathfrak{C}$ is a covering of S by elements of $\mathfrak{S}$. For every finite subcollection of $\mathfrak{C}$, the net (x_α) ultimately lies in the complement of the union of the members of the subcollection, so $\mathfrak{C}$ cannot be thinned to a finite subcovering of S. ■

It is instructive to compare this proof of Alexander's subbasis theorem to one based directly on the definition of compactness and some form of the axiom of choice, as can be found, for example, in [200, p. 368]. It is equally instructive to compare the following proof of Tychonoff's theorem to standard ones that directly apply the axiom of choice in one of its equivalent forms.

D.12 Tychonoff's Theorem. (A. N. Tychonoff, 1930 [234]; E. Čech, 1937 [40]). *Every topological product of compact topological spaces is compact.*

PROOF. Suppose that $\{ X^{(\alpha)} : \alpha \in I \}$ is a collection of compact topological spaces and that X is their topological product. It may be assumed that $I \neq \varnothing$. Let (x_β) be an ultranet in X. Since functions preserve ultranets, it follows that $(x_\beta^{(\alpha)})$ is an ultranet for each α in I, so there is an x in X such that $x_\beta^{(\alpha)} \to x^{(\alpha)}$ whenever $\alpha \in I$. Therefore $x_\beta \to x$, from which it follows that X is compact. ∎

References

MR citations refer to reviews published in *Mathematical Reviews*. Over the years, *Mathematical Reviews* has used three different citation systems. For example: MR **1**, 241 refers to a review on page 241 of volume 1; MR **21** #3766 refers to review #3766 in volume 21; and MR 81f:46001 refers to review #46001 in the sixth issue (issue f) of 1981.

[1] Naum I. Akhiezer and Mark G. Krein, *Some questions in the theory of moments*, Gosudarstv. Naučno-Tehn. Izdat. Ukrain., Kharkov, 1938 (Russian, translated into English as [2]).

[2] ———, *Some questions in the theory of moments*, Translations of Mathematical Monographs, vol. 2, American Mathematical Society, Providence, R.I., 1962 (English translation of [1]); MR **29** #5073.

[3] Leonidas Alaoglu, *Weak topologies of normed linear spaces*, Ann. of Math. (2) **41** (1940), 252–267; MR **1**, 241.

[4] James W. Alexander, *Ordered sets, complexes and the problem of compactification*, Proc. Nat. Acad. Sci. U.S.A. **25** (1939), 296–298.

[5] Kenneth W. Anderson, *Midpoint local uniform convexity, and other geometric properties of Banach spaces*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1960.

[6] Cesare Arzelà, *Funzioni di linee*, Atti della R. Accad. dei Lincei Rendiconti della Cl. Sci. Fis. Mat. Nat. (4) **5** [1° Sem.] (1889), 342–348.

[7] Giulio Ascoli, *Le curve limite di una varietà data di curve*, Atti della R. Accad. dei Lincei Memorie della Cl. Sci. Fis. Mat. Nat. (3) **18** (1882–1883), 521–586.

[8] Guido Ascoli, *Sugli spazi lineari metrici e le loro varietà lineari*, Ann. Mat. Pura Appl. (4) **10** (1932), 33–81, 203–232.

[9] Louis René Baire, *Sur les fonctions de variables réelles*, Ann. Mat. Pura Appl. (3) **3** (1899), 1–122.

[10] Stefan Banach, *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*, Fund. Math. **3** (1922), 133–181.

[11] ———, *Sur les fonctionelles linéaires*, Studia Math. **1** (1929), 211–216, 223–239.

[12] ———, *Über die Baire'sche Kategorie gewisser Funktionenmengen*, Studia Math. **3** (1931), 174–179.

[13] ———, *Théorie des opérations linéaires*, Monografje Matematyczne, Warsaw, 1932. See [15] for an English translation.

[14] ———, *Œuvres, Volume II*, PWN, Warsaw, 1979; MR 81f:46001. Includes a reprint of [13] along with comments by Aleksander Pełczyński and Czesław Bessaga summarizing much of the progress in Banach space theory from the appearance of [13] to 1979.

[15] ———, *Theory of linear operations*, North-Holland Mathematical Library, vol. 38, North-Holland, Amsterdam, 1987; MR 88a:01065. An English translation of [13] that includes the comments by Pełczyński and Bessaga from [14] on the progress in Banach space theory to 1979.

[16] Stefan Banach and Stanisław Mazur, *Zur Theorie der linearen Dimension*, Studia Math. **4** (1933), 100–112.

[17] Stefan Banach and Hugo Steinhaus, *Sur le principe de la condensation de singularités*, Fund. Math. **9** (1927), 50–61.

[18] Robert G. Bartle, *Nets and filters in topology*, Amer. Math. Monthly **62** (1955), 551–557; MR **17**, 391.

[19] ———, *The elements of real analysis*, second ed., Wiley, New York, 1976; MR **52** #14179.

[20] Robert G. Bartle, Nelson Dunford, and Jacob T. Schwartz, *Weak compactness and vector measures*, Canad. J. Math. **7** (1955), 289–305; MR **16**, 1123.

[21] Bernard Beauzamy, *Introduction to Banach spaces and their geometry*, Notas de Matemática [86], Mathematics Studies, vol. 68, North-Holland, Amsterdam, 1982; MR 84g:46017.

[22] Paul R. Beesack, Edward Hughes, and Marvin Ortel, *Rotund complex normed linear spaces*, Proc. Amer. Math. Soc. **75** (1979), 42–44; MR 81e:46013.

[23] Albert A. Bennett, *Newton's method in general analysis*, Proc. Nat. Acad. Sci. U.S.A. **2** (1916), 592–598.

[24] Sterling K. Berberian, *Introduction to Hilbert space*, second ed., Chelsea, New York, 1976; MR **25** #1424, MR **54** #3342.

[25] Czesław Bessaga and Aleksander Pełczyński, *A generalization of results of R. C. James concerning absolute bases in Banach spaces*, Studia Math. **17** (1958), 165–174; MR **22** #5874.

[26] ______, *On bases and unconditional convergence of series in Banach spaces*, Studia Math. **17** (1958), 151–164; MR **22** #5872.

[27] Garrett Birkhoff, *A note on topological groups*, Compositio Math. **3** (1936), 427–430.

[28] Errett Bishop and Robert R. Phelps, *A proof that every Banach space is subreflexive*, Bull. Amer. Math. Soc. **67** (1961), 97–98; MR **23** #A503.

[29] ______, *The support functionals of a convex set*, Convexity (Victor L. Klee, ed.), Proceedings of Symposia in Pure Mathematics, vol. 7, American Mathematical Society, Providence, R.I., 1963, pp. 27–35; MR **27** #4051.

[30] Jörg Blatter, *Reflexivity and the existence of best approximations*, Approximation Theory, II (George G. Lorentz, Charles K. Chui, and Larry L. Schumaker, eds.), Proc. Internat. Sympos., Univ. Texas, Austin, Tex., Academic Press, New York, 1976, pp. 299–301; MR **54** #13423.

[31] Ralph P. Boas, Jr., *A primer of real functions*, third ed., Carus Mathematical Monographs, vol. 13, Mathematical Association of America, Washington, D.C., 1981; MR **22** #9550, MR 83d:26002.

[32] Henri Frederic Bohnenblust and Andrew Sobczyk, *Extensions of functionals on complex linear spaces*, Bull. Amer. Math. Soc. **44** (1938), 91–93.

[33] Nicolas Bourbaki, *Éléments de mathématique*, Hermann et Cie., Paris, 1939 and later, various volumes; MR **3**, 55 is a review of the introductory volume.

[34] ———, *Éléments de mathématique, XV, Première partie: Les structures fondamentales de l'analyse, Livre V: Espaces vectoriels topologiques, Chapitres I et II*, Hermann et Cie., Act. Sci. et Ind. 1189, Paris, 1953; MR **14**, 880.

[35] Jean Bourgain, *On dentability and the Bishop-Phelps property*, Israel J. Math. **28** (1977), 265–271; MR **58** #2164.

[36] David G. Bourgin, *Linear topological spaces*, Amer. J. Math. **65** (1943), 637–659; MR **5**, 103.

[37] Richard D. Bourgin, *Geometric aspects of convex sets with the Radon-Nikodým property*, Lecture Notes in Mathematics, vol. 993, Springer-Verlag, Berlin, 1983; MR 85d:46023.

[38] Carl B. Boyer and Uta C. Merzbach, *A history of mathematics*, second ed., Wiley, New York, 1989; MR **38** #3105, MR 90f:01001.

[39] Peter G. Casazza, *"Classical sequences in Banach spaces," by Sylvie Guerre-Delabrière*, Bull. Amer. Math. Soc. (N.S.) **30** (1994), 117–124; book review.

[40] Eduard Čech, *On bicompact spaces*, Ann. of Math. (2) **38** (1937), 823–844.

[41] James A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc. **40** (1936), 396–414.

[42] John B. Conway, *Functions of one complex variable*, Graduate Texts in Mathematics, vol. 11, Springer-Verlag, New York, 1973; MR **56** #5843.

[43] ———, *A course in functional analysis*, Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1985; MR 86h:46001.

[44] Dennis F. Cudia, *The geometry of Banach spaces, smoothness*, Trans. Amer. Math. Soc. **110** (1964), 284–314; MR **29** #446.

[45] Alexander M. Davie, *The approximation problem for Banach spaces*, Bull. London Math. Soc. **5** (1973), 261–266; MR **49** #3499.

[46] Mahlon M. Day, *The spaces L^p with* $0 < p < 1$, Bull. Amer. Math. Soc. **46** (1940), 816–823; MR **2**, 102, MR **2**, 419.

[47] ———, *Reflexive Banach spaces not isomorphic to uniformly convex spaces*, Bull. Amer. Math. Soc. **47** (1941), 313–317; MR **2**, 221.

[48] ———, *Some more uniformly convex spaces*, Bull. Amer. Math. Soc. **47** (1941), 504–507; MR **2**, 314.

[49] ———, *Uniform convexity, III*, Bull. Amer. Math. Soc. **49** (1943), 745–750; MR **5**, 146.

[50] ———, *Uniform convexity in factor and conjugate spaces*, Ann. of Math. (2) **45** (1944), 375–385; MR **6**, 69.

[51] ———, *Strict convexity and smoothness of normed spaces*, Trans. Amer. Math. Soc. **78** (1955), 516–528; MR **16**, 716.

[52] ———, *Every L-space is isomorphic to a strictly convex space*, Proc. Amer. Math. Soc. **8** (1957), 415–417; MR **19**, 868.

[53] ———, *Normed linear spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 21, Springer-Verlag, Berlin, 1958; MR **20** #1187.

[54] ———, *Normed linear spaces*, second (corrected) ed., Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 21, Springer-Verlag, Berlin, 1962; MR **20** #1187, MR **26** #2847.

[55] ———, *On the basis problem in normed spaces*, Proc. Amer. Math. Soc. **13** (1962), 655–658; MR **25** #1435.

[56] ———, *Normed linear spaces*, third ed., Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 21, Springer-Verlag, Berlin, 1973; MR **20** #1187, MR **26** #2847, MR **49** #9588.

[57] Mahlon M. Day, Robert C. James, and Srinivasa Swaminathan, *Normed linear spaces that are uniformly convex in every direction*, Canad. J. Math. **23** (1971), 1051–1059; MR **44** #4492.

[58] Joseph Diestel, *Sequences and series in Banach spaces*, Graduate Texts in Mathematics, vol. 92, Springer-Verlag, New York, 1984; MR 85i:46020.

[59] Joseph Diestel and J. Jerry Uhl, *Vector measures*, Mathematical Surveys, vol. 15, American Mathematical Society, Providence, R.I., 1977; MR **56** #12216.

[60] Jean Dieudonné, *Sur le théorème de Hahn-Banach*, Revue Sci. (Rev. Rose Illus.) **79** (1941), 642–643; MR **7**, 124.

[61] ———, *Sur la séparation des ensembles convexes dans un espace de Banach*, Revue Sci. (Rev. Rose Illus.) **81** (1943), 277–278; MR **7**, 124.

[62] ———, *Sur la complétion des groupes topologiques*, C. R. Acad. Sci. Paris **218** (1944), 774–776; MR **7**, 241.

[63] ———, *Natural homomorphisms in Banach spaces*, Proc. Amer. Math. Soc. **1** (1950), 54–59; MR **11**, 524.

[64] ———, *History of functional analysis*, Notas de Matemática [77], Mathematics Studies, vol. 49, North-Holland, Amsterdam, 1981; MR 83d:46001.

[65] James Dugundji, *Topology*, Allyn and Bacon, Boston, 1966; MR **33** #1824.

[66] Nelson Dunford and Billy James Pettis, *Linear operations on summable functions*, Trans. Amer. Math. Soc. **47** (1940), 323–392; MR **1**, 338.

[67] Nelson Dunford and Jacob T. Schwartz, *Linear operators, part I: General theory*, Interscience, New York, 1958; MR **22** #8302.

[68] Peter L. Duren, *Theory of H^p spaces*, Academic Press, New York, 1970; MR **42** #3552.

[69] Peter L. Duren, Bernhard W. Romberg, and Allen L. Shields, *Linear functionals on H^p spaces with* $0 < p < 1$, J. Reine Angew. Math. **238** (1969), 32–60; MR **41** #4217.

[70] Aryeh Dvoretzky and Claude A. Rogers, *Absolute and unconditional convergence in normed linear spaces*, Proc. Nat. Acad. Sci. U.S.A. **36** (1950), 192–197; MR **11**, 525.

[71] William F. Eberlein, *Weak compactness in Banach spaces, I*, Proc. Nat. Acad. Sci. U.S.A. **33** (1947), 51–53; MR **9**, 42, MR **10**, 855.

[72] Nikolai V. Efimov and Sergei B. Stechkin, *Approximative compactness and Chebyshev sets*, Dokl. Akad. Nauk SSSR **140** (1961), 522–524 (Russian, translated into English as [73]); MR **25** #424.

[73] ———, *Approximative compactness and Chebyshev sets*, Soviet Math. Dokl. **2** (1961), 1226–1228 (English translation of [72]).

[74] M. Eidelheit, *Zur Theorie der konvexen Mengen in linearen normierten Räumen*, Studia Math. **6** (1936), 104–111.

[75] Per Enflo, *A Banach space with basis constant* > 1, Ark. Mat. **11** (1973), 103–107; MR **49** #7736.

[76] ———, *A counterexample to the approximation problem in Banach spaces*, Acta Math. **130** (1973), 309–317; MR **53** #6288.

[77] Ky Fan and Irving Glicksberg, *Some geometric properties of the spheres in a normed linear space*, Duke Math. J. **25** (1958), 553–568; MR **20** #5421.

[78] Tadeusz Figiel and William B. Johnson, *The approximation property does not imply the bounded approximation property*, Proc. Amer. Math. Soc. **41** (1973), 197–200; MR **49** #5782.

[79] Klaus Floret, *Weakly compact sets*, Lecture Notes in Mathematics, vol. 801, Springer-Verlag, Berlin, 1980; MR 82b:46001.

[80] Maurice René Fréchet, *Généralisation d'un théorème de Weierstrass*, C. R. Acad. Sci. Paris **139** (1904), 848–850.

[81] ———, *Sur quelques points du calcul fonctionnel*, Rend. Circ. Mat. Palermo **22** (1906), 1–74.

[82] Robert E. Fullerton, *Geometric structure of absolute basis systems in a linear topological space*, Pacific J. Math. **12** (1962), 137–147; MR **26** #5401.

[83] Vera Gantmacher, *Über schwache totalstetige Operatoren*, Mat. Sb. (N.S.) **7** (1940), 301–308; MR **2**, 224.

[84] Aleksandr L. Garkavi, *The best possible net and the best possible cross-section of a set in a normed space*, Izv. Akad. Nauk SSSR Ser. Mat. **26** (1962), 87–106 (Russian, translated into English as [85]); MR **25** #429.

[85] ———, *The best possible net and the best possible cross-section of a set in a normed space*, Amer. Math. Soc. Transl., Ser. 2 **39** (1964), 111–132 (English translation of [84]).

[86] Bernard R. Gelbaum, *Notes on Banach spaces and bases*, An. Acad. Brasil. Ciênc. **30** (1958), 29–36; MR **20** #5419.

[87] John R. Giles, *On smoothness of the Banach space embedding*, Bull. Austral. Math. Soc. **13** (1975), 69–74; MR **52** #1268.

[88] ———, *Convex analysis with application in the differentiation of convex functions*, Research Notes in Mathematics, vol. 58, Pitman, Boston, 1982; MR 83g:46001.

[89] Herman H. Goldstine, *Weakly complete Banach spaces*, Duke Math. J. **4** (1938), 125–131.

[90] W. Timothy Gowers, *Symmetric block bases of sequences with large average growth*, Israel J. Math. **69** (1990), 129–151; MR 91i:46010.

[91] ———, *A Banach space not containing* c_0, ℓ_1 *or a reflexive subspace*, Trans. Amer. Math. Soc. **344** (1994), 407–420; MR 94j:46024.

[92] ———, *A solution to Banach's hyperplane problem*, Bull. London Math. Soc. **26** (1994), 523–530; MR 96a:46025.

[93] ———, *A solution to the Schroeder-Bernstein problem for Banach spaces*, Bull. London Math. Soc. **28** (1996), 297–304; MR 97d:46009.

[94] W. Timothy Gowers and Bernard Maurey, *The unconditional basic sequence problem*, J. Amer. Math. Soc. **6** (1993), 851–874; MR 94k:46021.

[95] Alexandre Grothendieck, *Sur les applications linéaires faiblement compactes d'espaces du type* $C(K)$, Canad. J. Math. **5** (1953), 129–173; MR **15**, 438.

[96] ———, *Produits tensoriels topologiques et espaces nucléaires*, Mem. Amer. Math. Soc. **16** (1955), 1–140; MR **17**, 763.

[97] V. I. Gurariĭ, *The index of sequences in* $\tilde{C}$ *and the existence of infinite-dimensional separable Banach spaces having no orthogonal basis*, Rev. Roumaine Math. Pures Appl. **10** (1965), 967–971 (Russian); MR **34** #3282.

[98] Hans Hahn, *Über Folgen linearer Operationen*, Monatsh. Math. Phys. **32** (1922), 3–88.

[99] ———, *Über lineare Gleichungssysteme in linearen Räumen*, J. Reine Angew. Math. **157** (1927), 214–229.

[100] Paul R. Halmos, *"Nicolas Bourbaki,"* Sci. Amer. **196** (1957), no. 5, 88–99; MR **18**, 709.

[101] Felix Hausdorff, *Grundzüge der Mengenlehre*, Verlag von Veit, Leipzig, 1914.

[102] Eduard Helly, *Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten*, Monatsh. Math. Phys. **31** (1921), 60–91.

[103] David Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, IV*, Nachr. Kgl. Gesells. Wiss. Göttingen Math.-Phys. Kl. (1906), 157–227.

[104] Theophil Henry Hildebrandt, *On uniform limitedness of sets of functional operations*, Bull. Amer. Math. Soc. **29** (1923), 309–315.

[105] Kenneth Hoffman and Ray Kunze, *Linear algebra*, second ed., Prentice-Hall, Englewood Cliffs, N.J., 1971; MR **43** #1998.

[106] Thomas W. Hungerford, *Algebra*, Graduate Texts in Mathematics, vol. 73, Springer-Verlag, New York, 1980; MR **50** #6693, MR 82a:00006.

[107] Vasile I. Istrăţescu, *Strict convexity and complex strict convexity: Theory and applications*, Lecture Notes in Pure and Applied Mathematics, vol. 89, Marcel Dekker, New York, 1984; MR 86b:46023.

[108] Robert C. James, *Bases and reflexivity of Banach spaces*, Bull. Amer. Math. Soc. **56** (1950), 58, Abstract 80.

[109] ———, *Bases and reflexivity of Banach spaces*, Ann. of Math. (2) **52** (1950), 518–527; MR **12**, 616.

[110] ———, *A non-reflexive Banach space isometric with its second conjugate space*, Proc. Nat. Acad. Sci. U.S.A. **37** (1951), 174–177; MR **13**, 356.

[111] ———, *Reflexivity and the supremum of linear functionals*, Ann. of Math. (2) **66** (1957), 159–169; MR **19**, 755.

[112] ———, *Characterizations of reflexivity*, Studia Math. **23** (1964), 205–216; MR **30** #431.

[113] ———, *Uniformly non-square Banach spaces*, Ann. of Math. (2) **80** (1964), 542–550; MR **30** #4139.

[114] ———, *Weak compactness and reflexivity*, Israel J. Math. **2** (1964), 101–119; MR **31** #585.

[115] ———, *Weakly compact sets*, Trans. Amer. Math. Soc. **113** (1964), 129–140; MR **29** #2628.

[116] ———, *A counterexample for a sup theorem in normed spaces*, Israel J. Math. **9** (1971), 511–512; MR **43** #5287.

[117] ———, *Reflexivity and the sup of linear functionals*, Israel J. Math. **13** (1972), 289–300; MR **49** #3506.

[118] ———, *Bases in Banach spaces*, Amer. Math. Monthly **89** (1982), 625–640; MR 84h:46019.

[119] Mikhail I. Kadets, *On strong and weak convergence*, Dokl. Akad. Nauk SSSR **122** (1958), 13–16 (Russian); MR **20** #5422.

[120] ———, *Spaces isomorphic to a locally uniformly convex space*, Izv. Vyssh. Uchebn. Zaved. Mat. **1959** (1959), no. 6, 51–57 (Russian); MR **23** #A3987. See [121] for a correction to this paper.

[121] ———, *Letter to the editor*, Izv. Vyssh. Uchebn. Zaved. Mat. **1961** (1961), no. 6, 186–187 (Russian); MR **25** #3342. This contains a correction to [120].

[122] ———, *A proof of the topological equivalence of all separable infinite-dimensional Banach spaces*, Funktsional. Anal. i Prilozhen. **1** (1967), 61–70 (Russian, translated into English as [123]); MR **35** #700.

[123] ———, *A proof of the topological equivalence of all separable infinite-dimensional Banach spaces*, Functional Anal. Appl. **1** (1967), 53–62 (English translation of [122]).

[124] Shizuo Kakutani, *Über die Metrisation der topologischen Gruppen*, Proc. Imp. Acad. Tokyo **12** (1936), 82–84.

[125] ———, *Iteration of linear operations in complex Banach spaces*, Proc. Imp. Acad. Tokyo **14** (1938), 295–300.

[126] Nigel J. Kalton, *Basic sequences in F-spaces and their applications*, Proc. Edinburgh Math. Soc. (2) **19** (1974), 151–167; MR **54** #3350.

[127] ———, *The Orlicz-Pettis theorem*, Proceedings of the Conference on Integration, Topology, and Geometry in Linear Spaces held at the University of North Carolina, Chapel Hill, May 17–19, 1979 (William H. Graves, ed.), Contemporary Mathematics, vol. 2, American Mathematical Society, Providence, R.I., 1980, pp. 91–100; MR 82g:46027.

[128] Nigel J. Kalton, N. Tenney Peck, and James W. Roberts, *An F-space sampler*, London Mathematical Society Lecture Note Series, vol. 89, Cambridge University Press, Cambridge, 1984; MR 87c:46002.

[129] Irving Kaplansky, *Set theory and metric spaces*, Allyn and Bacon, Boston, 1972; MR **56** #11795.

[130] Samuel Karlin, *Bases in Banach spaces*, Duke Math. J. **15** (1948), 971–985; MR **10**, 548.

[131] John L. Kelley, *Convergence in topology*, Duke Math. J. **17** (1950), 277–283; MR **12**, 194.

[132] Victor L. Klee, *Some characterizations of reflexivity*, Rev. Cienc. (Lima) **52** (1950), 15–23; MR **13**, 250.

[133] ———, *Convex sets in linear spaces*, Duke Math. J. **18** (1951), 443–466; MR **13**, 354.

[134] ———, *Convex sets in linear spaces, II*, Duke Math. J. **18** (1951), 875–883; MR **13**, 849.

[135] ———, *Invariant metrics in groups (solution of a problem of Banach)*, Proc. Amer. Math. Soc. **3** (1952), 484–487; MR **13**, 848.

[136] ———, *Extremal structure of convex sets, II*, Math. Z. **69** (1958), 90–104; MR **19**, 1065.

[137] ———, *Some new results on smoothness and rotundity in normed linear spaces*, Math. Ann. **139** (1959), 51–63; MR **22** #5879.

[138] ———, *Mappings into normed linear spaces*, Fund. Math. **49** (1960/61), 25–34; MR **23** #A3985.

[139] ———, *A conjecture on weak compactness*, Trans. Amer. Math. Soc. **104** (1962), 398–402; MR **25** #3345.

[140] Konrad Knopp, *Infinite sequences and series*, Dover, New York, 1956; MR **18**, 30.

[141] Gottfried Köthe, *Topological vector spaces I*, Grundlehren der mathematischen Wissenschaften, vol. 159, Springer-Verlag, Berlin, 1983; revised second printing of volume reviewed earlier as MR **40** #1750.

[142] Vladimir Ya. Kozlov, *On bases in the space* $L_2[0,1]$, Mat. Sb. (N.S.) **26** (1950), 85–102 (Russian); MR **11**, 602.

[143] Mark G. Krein, *Sur quelques questions de la géométrie des ensembles convexes situés dans un espace linéaire normé et complet*, C. R. (Doklady) Acad. Sci. URSS (N.S.) **14** (1937), 5–7.

[144] Mark G. Krein, Mark A. Krasnoselskiĭ, and David P. Milman, *On the deficiency indices of linear operators in Banach spaces and some geometrical questions*, Sbornik Trudov Mat. Inst. Akad. Nauk Ukr. SSR **11** (1948), 97–112 (Russian).

[145] Mark G. Krein and David P. Milman, *On extreme points of regular convex sets*, Studia Math. **9** (1940), 133–137; MR **3**, 90.

[146] Mark G. Krein and Vitold L. Šmulian, *On regularly convex sets in the space conjugate to a Banach space*, Ann. of Math. (2) **41** (1940), 556–583; MR **1**, 335.

[147] Stanisław Kwapień, *On Banach spaces containing* c_0, Studia Math. **52** (1974), 187–188; MR **50** #8627.

[148] Ronald Larsen, *Banach algebras: An introduction*, Marcel Dekker, New York, 1973; MR **58** #7010.

[149] Henri Lebesgue, *Sur la divergence et la convergence non-uniforme des séries de Fourier*, C. R. Acad. Sci. Paris **141** (1905), 875–877.

[150] ———, *Leçons sur les séries trigonométriques*, Gauthier-Villars, Paris, 1906.

[151] ———, *Sur les intégrales singulières*, Ann. Fac. Sci. Univ. Toulouse (3) **1** (1909), 25–117.

[152] Jonathan Lewin, *A simple proof of Zorn's lemma*, Amer. Math. Monthly **98** (1991), 353–354.

[153] Joram Lindenstrauss, *On operators which attain their norm*, Israel J. Math. **1** (1963), 139–148; MR **28** #3308.

[154] ———, *On the modulus of smoothness and divergent series in Banach spaces*, Michigan Math. J. **10** (1963), 241–252; MR **29** #6316.

[155] Joram Lindenstrauss and Lior Tzafriri, *On the complemented subspaces problem*, Israel J. Math. **9** (1971), 263–269; MR **43** #2474.

[156] ———, *Classical Banach spaces I: Sequence spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 92, Springer-Verlag, Berlin, 1977; MR **58** #17766.

[157] ———, *Classical Banach spaces II: Function spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 97, Springer-Verlag, Berlin, 1979; MR 81c:46001.

[158] Edgar R. Lorch, *On a calculus of operators in reflexive vector spaces*, Trans. Amer. Math. Soc. **45** (1939), 217–234.

[159] A. R. Lovaglia, *Locally uniformly convex Banach spaces*, Trans. Amer. Math. Soc. **78** (1955), 225–238; MR **16**, 596.

[160] R. Daniel Mauldin (ed.), *The Scottish book: Mathematics from the Scottish Café*, Birkhäuser, Boston, 1981; MR 84m:00015.

[161] Stanisław Mazur, *Über die kleinste konvexe Menge, die eine gegebene kompakte Menge enthält*, Studia Math. **2** (1930), 7–9.

[162] ———, *Über konvexe Mengen in linearen normierten Räumen*, Studia Math. **4** (1933), 70–84.

[163] ———, *Über schwache Konvergenz in den Räumen (L^p)*, Studia Math. **4** (1933), 128–133.

[164] Stanisław Mazur and Władysław Orlicz, *Über Folgen linearer Operationen*, Studia Math. **4** (1933), 152–157.

[165] Stanisław Mazur and L. Sternbach, *Über die Borelschen Typen von linearen Mengen*, Studia Math. **4** (1933), 48–53.

[166] Edward J. McShane, *Linear functionals on certain Banach spaces*, Proc. Amer. Math. Soc. **1** (1950), 402–408; MR **12**, 10.

[167] ______, *A theory of limits*, Studies in Modern Analysis (R. Creighton Buck, ed.), MAA Studies in Mathematics, vol. 1, Mathematical Association of America, Washington, D.C., 1962, pp. 7–29.

[168] Robert E. Megginson, *The semi-Kadec-Klee condition and nearest-point properties of sets in normed linear spaces*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1984.

[169] David P. Milman, *On some criteria for the regularity of spaces of the type* (B), C. R. (Doklady) Acad. Sci. URSS (N.S.) **20** (1938), 243–246.

[170] ______, *Characteristics of extremal points of regularly convex sets*, Dokl. Akad. Nauk SSSR **57** (1947), 119–122 (Russian); MR **9**, 192.

[171] Eliakim H. Moore and Henry L. Smith, *A general theory of limits*, Amer. J. Math. **44** (1922), 102–121.

[172] James R. Munkres, *Topology: A first course*, Prentice-Hall, Englewood Cliffs, N.J., 1975; MR **57** #4063.

[173] Masahiro Nakamura, *Complete continuities of linear operators*, Proc. Japan Acad. **27** (1951), 544–547; MR **14**, 183.

[174] Masahiro Nakamura and Shizuo Kakutani, *Banach limits and the Čech compactification of a countable discrete set*, Proc. Imp. Acad. Tokyo **19** (1943), 224–229; MR **7**, 306.

[175] Otton M. Nikodým, *Contribution à la théorie des fonctionnelles linéaires en connexion avec la théorie de la mesure des ensembles abstraits*, Mathematica (Cluj) **5** (1931), 130–141.

[176] Göte Nordlander, *The modulus of convexity in normed linear spaces*, Ark. Mat. **4** (1960), 15–17; MR **25** #4329.

[177] Władysław Orlicz, *Beiträge zur Theorie der Orthogonalentwicklungen II*, Studia Math. **1** (1929), 241–255.

[178] John C. Oxtoby, *Measure and category*, second ed., Graduate Texts in Mathematics, vol. 2, Springer-Verlag, New York, 1980; MR **52** #14213, MR 81j:28003.

[179] Aleksander Pełczyński, *A note on the paper of I. Singer "Basic sequences and reflexivity of Banach spaces,"* Studia Math. **21** (1962), 371–374; MR **26** #4156.

[180] Aleksander Pełczyński and Vladimir N. Sudakov, *Remark on non-complemented subspaces of the space* $m(S)$, Colloq. Math. **9** (1962), 85–88; MR **25** #2441.

[181] Billy James Pettis, *A note on regular Banach spaces*, Bull. Amer. Math. Soc. **44** (1938), 420–428.

[182] ———, *On integration in vector spaces*, Trans. Amer. Math. Soc. **44** (1938), 277–304.

[183] ———, *A proof that every uniformly convex space is reflexive*, Duke Math. J. **5** (1939), 249–253.

[184] Robert R. Phelps, *Subreflexive normed linear spaces*, Arch. Math. (Basel) **8** (1957), 444–450; MR **20** #6027. See [185] for a correction to this paper.

[185] ———, *Correction to "Subreflexive normed linear spaces,"* Arch. Math. (Basel) **9** (1958), 439–440; MR **21** #2175. This contains a correction to [184].

[186] ———, *A representation theorem for bounded convex sets*, Proc. Amer. Math. Soc. **11** (1960), 976–983; MR **23** #A501.

[187] ———, *The Bishop-Phelps theorem in complex spaces: An open problem*, Function Spaces (Krzysztof Jarosz, ed.), Lecture Notes in Pure and Applied Mathematics, vol. 136, Marcel Dekker, New York, 1992, pp. 337–340; MR 93e:46022.

[188] Ralph S. Phillips, *Integration in a convex linear topological space*, Trans. Amer. Math. Soc. **47** (1940), 114–145; MR **2**, 103.

[189] ———, *On linear transformations*, Trans. Amer. Math. Soc. **48** (1940), 516–541; MR **2**, 318.

[190] Mauro Picone, *Lezioni di analisi infinitesimale*, vol. 1, Circolo Matematico di Catania, Catania, Italy, 1923; preface reproduced in [191].

[191] ———, *Lezioni di analisi infinitesimale: Prefazione al vol. I*, Esercitazioni Matematiche: Pubblicazione del Circolo Matematico di Catania **3** (1923), 118–122.

[192] Johann Radon, *Theorie und Anwendungen der absolut additiven Mengenfunktionen*, S.-B. Akad. Wiss. Wien **122** (1913), 1295–1438.

[193] Frédéric Riesz, *Untersuchungen über Systeme integrierbarer Funktionen*, Math. Ann. **69** (1910), 449–497.

[194] ———, *Les systèmes d'équations linéaires à une infinité d'inconnues*, Gauthier-Villars, Paris, 1913.

[195] ———, *Über lineare Funktionalgleichungen*, Acta Math. **41** (1918), 71–98.

[196] ———, *Sur la convergence en moyenne, I*, Acta Sci. Math. (Szeged) **4** (1928–1929), 58–64.

[197] ———, *Sur la convergence en moyenne, II*, Acta Sci. Math. (Szeged) **4** (1928–1929), 182–185.

[198] Ira Rosenholtz, *Another proof that any compact metric space is the continuous image of the Cantor set*, Amer. Math. Monthly **83** (1976), 646–647; MR **54** #6064.

[199] Haskell P. Rosenthal, *A characterization of Banach spaces containing ℓ_1*, Proc. Nat. Acad. Sci. U.S.A. **71** (1974), 2411–2413; MR **50** #10773.

[200] Walter Rudin, *Functional analysis*, McGraw-Hill, New York, 1973; MR **51** #1315.

[201] ———, *Principles of mathematical analysis*, third ed., McGraw-Hill, New York, 1976; MR **52** #5893.

[202] ———, *Real and complex analysis*, third ed., McGraw-Hill, New York, 1987; MR 88k:00002.

[203] Anthony F. Ruston, *A note on convexity in Banach spaces*, Proc. Cambridge Philos. Soc. **45** (1949), 157–159; MR **10**, 197.

[204] Donald Sarason, *The exact answer to a question of Shields*, Math. Intelligencer **12** (1990), no. 2, 18–19; MR 91d:01006h.

[205] Helmut H. Schaefer, *Topological vector spaces*, Macmillan, New York, 1966; MR **33** #1689.

[206] Juliusz Schauder, *Zur Theorie stetiger Abbildungen in Funktionalräumen*, Math. Z. **26** (1927), 47–65.

[207] ———, *Eine Eigenschaft des Haarschen Orthogonalsystems*, Math. Z. **28** (1928), 317–320.

[208] ———, *Über die Umkehrung linearer, stetiger Funktionaloperationen*, Studia Math. **2** (1930), 1–6.

[209] ———, *Über lineare, vollstetige Funktionaloperationen*, Studia Math. **2** (1930), 183–196.

[210] Thomas Schlumprecht, *An arbitrarily distortable Banach space*, Israel J. Math. **76** (1991), 81–95; MR 93h:46023.

[211] Erhard Schmidt, *Über die Auflösung linearer Gleichungen mit unendlich vielen Unbekannten*, Rend. Circ. Mat. Palermo **25** (1908), 53–77.

[212] J. Schur, *Über lineare Transformationen in der Theorie der unendlichen Reihen*, J. Reine Angew. Math. **151** (1920), 79–111.

[213] Wacław Sierpiński, *Sur les ensembles complets d'un espace* (*D*), Fund. Math. **11** (1928), 203–205.

[214] ———, *Cardinal and ordinal numbers*, Polska Akademia Nauk, Monografie Matematyczne, vol. 34, Państwowe Wydawnictwo Naukowe, Warsaw, 1958; MR **20** #2288.

[215] Ivan Singer, *Some remarks on approximative compactness*, Rev. Roumaine Math. Pures Appl. **9** (1964), 167–177; MR **31** #2707.

[216] ———, *Bases in Banach spaces I*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, vol. 154, Springer-Verlag, Berlin, 1970; MR **45** #7451.

[217] Mark A. Smith, *Directional uniform rotundity in Banach spaces*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1975.

[218] ———, *Banach spaces that are uniformly rotund in weakly compact sets of directions*, Canad. J. Math. **29** (1977), 963–970; MR **56** #9232.

[219] ———, *Some examples concerning rotundity in Banach spaces*, Math. Ann. **233** (1978), 155–161; MR **58** #2168.

[220] ———, *A Banach space that is MLUR but not HR*, Math. Ann. **256** (1981), 277–279; MR 82h:46032.

[221] Vitold L. Šmulian, *On some geometrical properties of the sphere in a space of the type* (*B*), C. R. (Doklady) Acad. Sci. URSS (N.S.) **24** (1939), 648–652; MR **1**, 242.

[222] ———, *On some geometrical properties of the unit sphere in the space of the type* (*B*), Mat. Sb. (N.S.) **6** (1939), 77–94 (Russian, English summary); MR **1**, 242.

[223] ———, *On the principle of inclusion in the space of the type (B)*, Mat. Sb. (N.S.) **5** (1939), 317–328 (Russian, English summary); MR **1**, 335.

[224] ———, *Sur la dérivabilité de la norme dans l'espace de Banach*, C. R. (Doklady) Acad. Sci. URSS (N.S.) **27** (1940), 643–648; MR **2**, 102.

[225] ———, *Über lineare topologische Räume*, Mat. Sb. (N.S.) **7** (1940), 425–448 (German, Russian summary); MR **2**, 102.

[226] ———, *Sur la structure de la sphère unitaire dans l'espace de Banach*, Mat. Sb. (N.S.) **9** (1941), 545–561; MR **3**, 205.

[227] G. A. Soukhomlinoff, *Über Fortsetzung von linearen Funktionalen in linearen komplexen Räumen und linearen Quaternionräumen*, Mat. Sb. (N.S.) **3** (1938), 353–358 (Russian, German summary).

[228] Lynn A. Steen and J. Arthur Seebach, Jr., *Counterexamples in topology*, Holt, Rinehart and Winston, New York, 1970; MR **42** #1040.

[229] Stefan Straszewicz, *Über exponierte Punkte abgeschlossener Punktmengen*, Fund. Math. **24** (1935), 139–143.

[230] Stanisław J. Szarek, *A Banach space without a basis which has the bounded approximation property*, Acta Math. **159** (1987), 81–98; MR 88f:46029.

[231] Stanimir L. Ţroyanski, *An example of a smooth space whose dual is not strictly normed*, Studia Math. **35** (1970), 305–309 (Russian); MR **42** #6589.

[232] ———, *On locally uniformly convex and differentiable norms in certain non-separable Banach spaces*, Studia Math. **37** (1971), 173–180; MR **46** #5995.

[233] John W. Tukey, *Some notes on the separation of convex sets*, Portugal. Math. **3** (1942), 95–102; MR **4**, 13.

[234] Andrei N. Tychonoff, *Über die topologische Erweiterung von Räumen*, Math. Ann. **102** (1930), 544–561.

[235] J. Jerry Uhl, *Pettis's measurability theorem*, Proceedings of the Conference on Integration, Topology, and Geometry in Linear Spaces held at the University of North Carolina, Chapel Hill, May 17–19, 1979 (William H. Graves, ed.), Contemporary Mathematics, vol. 2, American Mathematical Society, Providence, R.I., 1980, pp. 135–144; MR 82m:28020.

[236] Leonid P. Vlasov, *Some theorems on Chebyshev sets*, Mat. Zametki **11** (1972), 135–144 (Russian, translated into English as [237]); MR **45** #9046.

[237] ———, *Some theorems on Chebyshev sets*, Math. Notes **11** (1972), 87–92 (English translation of [236]).

[238] ———, *Approximative properties of sets in normed linear spaces*, Uspekhi Mat. Nauk **28** (1973), 3–66 (Russian, translated into English as [239]); MR **53** #8761.

[239] ———, *Approximative properties of sets in normed linear spaces*, Russian Math. Surveys **28** (1973), 1–66 (English translation of [238]).

[240] Rudolf Vyborny, *On the weak convergence in locally uniformly convex spaces*, Časopis Pěst. Mat. **81** (1956), 352–353 (Czech).

[241] John V. Wehausen, *Transformations in linear topological spaces*, Duke Math. J. **4** (1938), 157–169.

[242] Richard L. Wheeden and Antoni Zygmund, *Measure and integral: An introduction to real analysis*, Marcel Dekker, New York, 1977; MR **58** #11295.

[243] Robert F. Wheeler, *The equicontinuous weak* topology and semi-reflexivity*, Studia Math. **41** (1972), 243–256; MR **46** #4150.

[244] Robert J. Whitley, *Projecting m onto* c_0, Amer. Math. Monthly **73** (1966), 285–286.

[245] Kôsaku Yosida, *Mean ergodic theorem in Banach spaces*, Proc. Imp. Acad. Tokyo **14** (1938), 292–294.

[246] Petr P. Zabreĭko, *A theorem for semiadditive functionals*, Funktsional. Anal. i Prilozhen. **3** (1969), 86–88 (Russian, translated into English as [247]); MR **39** #3283.

[247] ———, *A theorem for semiadditive functionals*, Functional Anal. Appl. **3** (1969), 70–72 (English translation of [246]).

[248] Vaclav E. Zizler, *On some rotundity and smoothness properties of Banach spaces*, Dissertationes Math. (Rozprawy Mat.) **87** (1971), 33 pp. + errata insert; MR **45** #9108.

[249] Max A. Zorn, *A remark on method in transfinite algebra*, Bull. Amer. Math. Soc. **41** (1935), 667–670.

List of Symbols

Fraktur Letters

Here are the uppercase Roman letters, each followed by its Fraktur equivalent:

A, $\mathfrak{A}$, B, $\mathfrak{B}$, C, $\mathfrak{C}$, D, $\mathfrak{D}$, E, $\mathfrak{E}$, F, $\mathfrak{F}$, G, $\mathfrak{G}$, H, $\mathfrak{H}$, I, $\mathfrak{I}$, J, $\mathfrak{J}$, K, $\mathfrak{K}$, L, $\mathfrak{L}$, M, $\mathfrak{M}$, N, $\mathfrak{N}$, O, $\mathfrak{O}$, P, $\mathfrak{P}$, Q, $\mathfrak{Q}$, R, $\mathfrak{R}$, S, $\mathfrak{S}$, T, $\mathfrak{T}$, U, $\mathfrak{U}$, V, $\mathfrak{V}$, W, $\mathfrak{W}$, X, $\mathfrak{X}$, Y, $\mathfrak{Y}$, Z, $\mathfrak{Z}$.

Other Symbols

Symbols used as the names of particular spaces, such as c_0, have in most cases been placed in the Index rather than here. Also, most symbols invented for temporary use, such as those in the discussion of arrow vectors at the beginning of Section 1.2, have been omitted from this list.

Blackboard Bold Symbols

Symbols for Norms

Symbols for Dual Spaces and Normed Spaces Formed from Other Normed Spaces

Symbols for Properties of Normed Spaces

Symbols for Moduli of Rotundity and Smoothness

Symbols for Spaces of Linear Operators

Other Symbols Involving Linear Operators

Symbols Used as Technical Notation for Algebraic, Analytic, and Topological Systems

Symbols for Convergence, Limits, and Sums

Symbols for Metrics and Other Distance Functions

Symbols for Sets Formed from Other Sets

Symbols for Binary Relations and Unary and Binary Operations

Miscellaneous Symbols

Index

References to the basic material on topology and metric spaces in Section 2.1 and Appendices B and D have been put in separate subindices following the main index. For example, references to the basic properties of open sets given in these background sections can be found under "metric space theory, open set" and "topological theory, open set" in these subindices, while references outside these sections to properties of open sets in settings more specific to the subject of this book appear under the separate entry for "open set" in the main index.

Graduate Texts in Mathematics

(continued from page ii)

120 ZIEMER. Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation.
121 LANG. Cyclotomic Fields I and II. Combined 2nd ed.
122 REMMERT. Theory of Complex Functions. *Readings in Mathematics*
123 EBBINGHAUS/HERMES et al. Numbers. *Readings in Mathematics*
124 DUBROVIN/FOMENKO/NOVIKOV. Modern Geometry—Methods and Applications. Part III.
125 BERENSTEIN/GAY. Complex Variables: An Introduction.
126 BOREL. Linear Algebraic Groups. 2nd ed.
127 MASSEY. A Basic Course in Algebraic Topology.
128 RAUCH. Partial Differential Equations.
129 FULTON/HARRIS. Representation Theory: A First Course. *Readings in Mathematics*
130 DODSON/POSTON. Tensor Geometry.
131 LAM. A First Course in Noncommutative Rings.
132 BEARDON. Iteration of Rational Functions.
133 HARRIS. Algebraic Geometry: A First Course.
134 ROMAN. Coding and Information Theory.
135 ROMAN. Advanced Linear Algebra.
136 ADKINS/WEINTRAUB. Algebra: An Approach via Module Theory.
137 AXLER/BOURDON/RAMEY. Harmonic Function Theory.
138 COHEN. A Course in Computational Algebraic Number Theory.
139 BREDON. Topology and Geometry.
140 AUBIN. Optima and Equilibria. An Introduction to Nonlinear Analysis.
141 BECKER/WEISPFENNING/KREDEL. Gröbner Bases. A Computational Approach to Commutative Algebra.
142 LANG. Real and Functional Analysis. 3rd ed.
143 DOOB. Measure Theory.
144 DENNIS/FARB. Noncommutative Algebra.
145 VICK. Homology Theory. An Introduction to Algebraic Topology. 2nd ed.
146 BRIDGES. Computability: A Mathematical Sketchbook.
147 ROSENBERG. Algebraic K-Theory and Its Applications.
148 ROTMAN. An Introduction to the Theory of Groups. 4th ed.
149 RATCLIFFE. Foundations of Hyperbolic Manifolds.
150 EISENBUD. Commutative Algebra with a View Toward Algebraic Geometry.
151 SILVERMAN. Advanced Topics in the Arithmetic of Elliptic Curves.
152 ZIEGLER. Lectures on Polytopes.
153 FULTON. Algebraic Topology: A First Course.
154 BROWN/PEARCY. An Introduction to Analysis.
155 KASSEL. Quantum Groups.
156 KECHRIS. Classical Descriptive Set Theory.
157 MALLIAVIN. Integration and Probability.
158 ROMAN. Field Theory.
159 CONWAY. Functions of One Complex Variable II.
160 LANG. Differential and Riemannian Manifolds.
161 BORWEIN/ERDÉLYI. Polynomials and Polynomial Inequalities.
162 ALPERIN/BELL. Groups and Representations.
163 DIXON/MORTIMER. Permutation Groups.
164 NATHANSON. Additive Number Theory: The Classical Bases.
165 NATHANSON. Additive Number Theory: Inverse Problems and the Geometry of Sumsets.
166 SHARPE. Differential Geometry: Cartan's Generalization of Klein's Erlangen Program.
167 MORANDI. Field and Galois Theory.
168 EWALD. Combinatorial Convexity and Algebraic Geometry.
169 BHATIA. Matrix Analysis.
170 BREDON. Sheaf Theory. 2nd ed.
171 PETERSEN. Riemannian Geometry.
172 REMMERT. Classical Topics in Complex Function Theory.
173 DIESTEL. Graph Theory.
174 BRIDGES. Foundations of Real and Abstract Analysis.
175 LICKORISH. An Introduction to Knot Theory.
176 LEE. Riemannian Manifolds.
177 NEWMAN. Analytic Number Theory.
178 CLARKE/LEDYAEV/STERN/WOLENSKI. Nonsmooth Analysis and Control Theory.
179 DOUGLAS. Banach Algebra Techniques in Operator Theory. 2nd ed.

180 SRIVASTAVA. A Course on Borel Sets.
181 KRESS. Numerical Analysis.
182 WALTER. Ordinary Differential Equations.
183 MEGGINSON. An Introduction to Banach Space Theory.
184 BOLLOBAS. Modern Graph Theory.
185 COX/LITTLE/O'SHEA. Using Algebraic Geometry.
186 RAMAKRISHNAN/VALENZA. Fourier Analysis on Number Fields.
187 HARRIS/MORRISON. Moduli of Curves.
188 GOLDBLATT. Lectures on the Hyperreals: An Introduction to Nonstandard Analysis.
189 LAM. Lectures on Modules and Rings.

书　　名： An Introduction to Banach Space Theory

作　　者： R. E. Megginson

中 译 名： 巴拿赫空间理论引论

出 版 者： 世界图书出版公司北京公司

印 刷 者： 北京世图印刷厂

发　　行： 世界图书出版公司北京公司（北京朝内大街 137 号　100010）

联系电话： 010-64015659, 64038347

电子信箱： kjsk@vip.sina.com

开　　本： 24　　　印　张：26

出版年代： 2003 年 6 月

书　　号： 7-5062-5964-8 / O・383

版权登记： 图字:01-2003-3765

定　　价： 58.00 元