Universitext

Douglas S. Bridges
Luminiţa Simona Vîţă

Techniques of Constructive Analysis

Springer

Universitext

Universitext

Editors (North America): S. Axler and K.A. Ribet

Aguilar/Gitler/Prieto: Algebraic Topology from a Homotopical Viewpoint
Aksoy/Khamsi: Nonstandard Methods in Fixed Point Theory
Andersson: Topics in Complex Analysis
Aupetit: A Primer on Spectral Theory
Bachman/Narici/Beckenstein: Fourier and Wavelet Analysis
Badescu: Algebraic Surfaces
Balakrishnan/Ranganathan: A Textbook of Graph Theory
Balser: Formal Power Series and Linear Systems of Meromorphic Ordinary Differential Equations
Bapat: Linear Algebra and Linear Models (2nd ed.)
Berberian: Fundamentals of Real Analysis
Blyth: Lattices and Ordered Algebraic Structures
Boltyanskii/Efremovich: Intuitive Combinatorial Topology. (Shenitzer, trans.)
Booss/Bleecker: Topology and Analysis
Borkar: Probability Theory: An Advanced Course
Böttcher/Silbermann: Introduction to Large Truncated Toeplitz Matrices
Bridges/Vîţă: Techniques of Constructive Analysis
Carleson/Gamelin: Complex Dynamics
Cecil: Lie Sphere Geometry: With Applications to Submanifolds
Chae: Lebesgue Integration (2nd ed.)
Charlap: Bieberbach Groups and Flat Manifolds
Chern: Complex Manifolds Without Potential Theory
Cohn: A Classical Invitation to Algebraic Numbers and Class Fields
Curtis: Abstract Linear Algebra
Curtis: Matrix Groups
Debarre: Higher-Dimensional Algebraic Geometry
Deitmar: A First Course in Harmonic Analysis (2nd ed.)
DiBenedetto: Degenerate Parabolic Equations
Dimca: Singularities and Topology of Hypersurfaces
Edwards: A Formal Background to Mathematics I a/b
Edwards: A Formal Background to Mathematics II a/b
Engel/Nagel: A Short Course on Operator Semigroups
Farenick: Algebras of Linear Transformations
Foulds: Graph Theory Applications
Friedman: Algebraic Surfaces and Holomorphic Vector Bundles
Fuhrmann: A Polynomial Approach to Linear Algebra
Gardiner: A First Course in Group Theory
Gårding/Tambour: Algebra for Computer Science
Goldblatt: Orthogonality and Spacetime Geometry
Gustafson/Rao: Numerical Range: The Field of Values of Linear Operators and Matrices
Hahn: Quadratic Algebras, Clifford Algebras, and Arithmetic Witt Groups
Heinonen: Lectures on Analysis on Metric Spaces
Holmgren: A First Course in Discrete Dynamical Systems
Howe/Tan: Non-Abelian Harmonic Analysis: Applications of $SL(2, \mathrm{R})$
Howes: Modern Analysis and Topology
Hsieh/Sibuya: Basic Theory of Ordinary Differential Equations
Humi/Miller: Second Course in Ordinary Differential Equations

(continued after index)

Douglas S. Bridges and Luminiţa Simona Vîţă

Techniques of Constructive Analysis

Springer

Douglas S. Bridges
Department of Mathematics/Statistics
University of Canterbury
Christchurch, New Zealand
d.bridges@math.canterbury.ac.nz

Luminiţa Simona Vîţă
Department of Mathematics/Statistics
University of Canterbury
Christchurch, New Zealand
l.vita@math.canterbury.ac.nz

Mathematics Subject Classification (2000): 03F60, 26E40, 46S30, 47S30, 03F55, 03F65, 68Q99

Library of Congress Control Number: 2006926441

ISBN-10: 0-387-33646-X e-ISBN-10: 0-387-38147-3
ISBN-13: 978-0387-33646-6 e-ISBN-13: 978-0387-38147-3

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

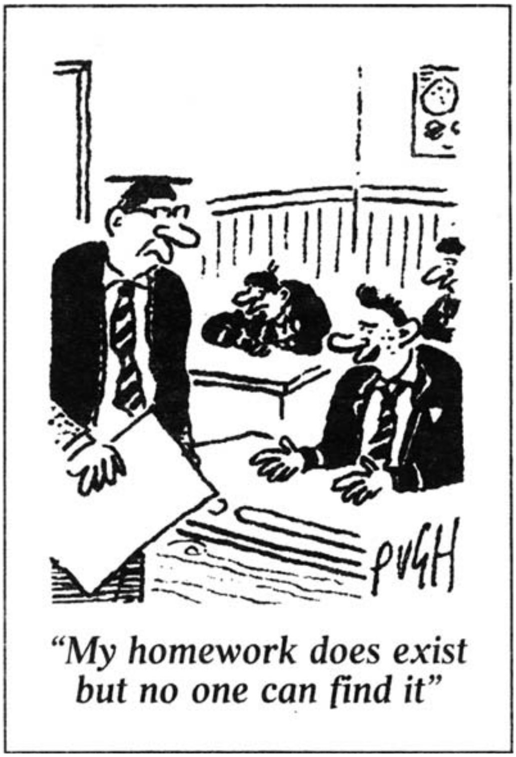
springer.com

Ideal pierdut în noaptea unei lumi ce nu mai este,
Lume ce gândea în basme şi vorbea în poezii,
O! te văd, te-aud, te cuget, tânără şi dulce veste
Dintr-un cer cu alte stele, cu-alte raiuri, cu alţi zei.

—Mihai Eminescu, "Venere şi Madonă"

Oh, ideal lost in night-mists of a vanished universe:
People who would think in legends—all a world who spoke in verse;
I can see and think and hear you—youthful scout which gently nods
From a sky with different starlights, other Edens, other gods.

—Mihai Eminescu, "Venus and Madonna"
(translated by Andrei Bantaş)



This image is a courtesy of The Times of London. Printed in the February 3, 2004 issue.

Preface

ROSENCRANTZ: *Shouldn't we be doing something... constructive?*

GUILDENSTERN*: What did you have in mind?*

—Tom Stoppard, *Rosencratz and Guildenstern Are Dead*

We have written this book in order to provide an introduction to constructive analysis, emphasising techniques and results that have been obtained in the last twenty years. The intended readership comprises senior undergraduates, postgraduates, and professional researchers in mathematics and theoretical computer science. We hope that our work will help spread the message that doing mathematics constructively is interesting (it can even be fun!) and challenging, and produces new, deep computational information.

An appreciation of the distinction between constructive and nonconstructive has become more widespread in this era of computers. Nevertheless, there are few books devoted to the development of mathematics in a rigorously constructive/computable fashion, although there are some, primarily concentrating on logic and foundations, in which the odd chapter deals with constructive mathematics proper as distinct from its underlying logic or set theory. It is now almost forty years since the publication of Errett Bishop's seminal monograph *Foundations of Constructive Analysis* [9], which in our view is one of the most remarkable intellectual documents of the twentieth century, and more than twenty since the appearance of its outgrowth [12]. In the intervening years there has been considerable activity in constructive analysis, algebra, and topology; in related foundational areas such as type theory [69]; and in the relation between constructive mathematics and computer science (for example, program extraction from proofs [42, 70, 51]). Believing that a new introduction to the mathematical, as distinct from the foundational, side of the subject is overdue, we embarked upon this monograph.

Our book is intended not to replace, but to supplement, Bishop's original classic [9] and the later volume [12] based thereon. Both of those two monographs cover aspects of analysis, such as Haar measure and commutative Banach algebras, that we do not mention. We cover some topics that are found in [9] and [12] (it would be almost inconceivable to produce a book like ours, dealing with constructive mathematics for nonexperts, without proving, for example, basic results about locatedness and total boundedness); but we have tried to provide improved proofs whenever possible. However, much of the material we present was simply not around at the time of writing of [9] or [12].

Instead of systematically developing analysis, beginning with the real line and continuing through metric, normed, and Hilbert spaces to its higher reaches, we have chosen to write the chapters around certain themes or techniques (hence our title). For example, Chapter 3 is devoted to the λ-technique, which, since its first use in the proof of Lemma 7 on page 177 of [9], has become a surprisingly powerful tool with applications in many areas of constructive analysis. A major influence in the application of the λ-technique was Ishihara's remarkable paper [60], which showed that a subtle use of the technique could enable us to prove disjunctions whose proof, although trivial with classical logic, appears at first sight to be constructively out of the question. This paper opened up many new pathways in constructive analysis.

Chapter 1 introduces constructive mathematics and lays the foundations for the later chapters. In Chapter 2 we first present a new construction of the real numbers, motivated by ideas in [2]. After deriving standard properties such as the completeness of $\mathbb{R}$, we introduce metric spaces, with the major theme of locatedness, and normed linear spaces. When we discuss metric, normed, and Hilbert spaces, we assume some familiarity with the standard classical definitions of those concepts and with those elementary classical properties that pass over unchanged to the constructive setting.

Chapter 3 we have already referred to. The main theme of Chapter 4 is finite-dimensionality, but the chapter concludes with an introduction to Hilbert spaces.

Chapter 5 deals with convexity in normed spaces. Starting with some elementary convex geometry in $\mathbb{R}^n$, the chapter goes on to handle separation and Hahn–Banach theorems, locally convex spaces, and duality. Following Bishop, we describe those linear functionals that are weak*-uniformly continuous on the unit ball of the dual space. We then give a new application of the technique used to prove that result, thereby characterising certain continuous linear functionals on the space of bounded operators on a Hilbert space.

In Chapter 6 we derive a range of results associated with the theme of locatedness and with the λ-technique introduced in Chapter 3. We pay particular attention to necessary and sufficient conditions for convex subsets of a normed space to be located, and to connections between properties of an operator on a Hilbert space and those of its adjoint—when that adjoint exists: it may not always do so constructively. The final section of the book deals with a relatively recent version of Baire's theorem and its applications, and culminates in constructive versions of three of the big guns in functional analysis: the open mapping, inverse mapping, and closed graph theorems.

Which parts of the book deal with new material, compared with what appeared in [12]? We have already mentioned the new construction of the real numbers, in Chapter 2. Notable novelties in the later chapters include all but one result in Chapter 3 on the λ-technique; the section on convexity, Ishihara's results on exact Hahn–Banach extensions, and our characterisation theorem for certain continuous linear functionals, all in Chapter 5; and virtually all of Chapter 6. Throughout the book there are what we hope will be seen as improvements and simplifications of proofs of many results that were given in [9] or [12].

What do we mean by "constructive analysis" in the title of this book? We do *not* mean analysis carried out with the usual "classical" logic within a framework, such as recursive function theory, designed to capture the concept of computability. In our view, such a notion of constructive has at least two drawbacks. First, by working within, say, the recursive setting, it can make the mathematics look less like normal mathematics and much harder to read. Secondly, the recursive constraint removes the possibility of other interpretations of the mathematics, such as Brouwer's intuitionistic one [48]. Our approach, on the other hand, has neither of these features: the mathematics looks and reads just like the mathematics one is used to from undergraduate days, and all our proofs and results are valid in several models. They are valid in the recursive model, in intuitionistic mathematics, and, we believe, in any of the models for "computable mathematics" (including Weihrauch's Type Two Effectivity Theory [91], within which Andrej Bauer has recently found a realisability interpretation of constructive mathematics within Weihrauch's theory [5]). *They are also valid proofs in standard mathematics with classical logic.* For example, our proof of the Hahn–Banach theorem (Theorem 5.3.3) is, as it stands, a valid algorithmic proof of the classical Hahn–Banach theorem. Moreover—and this is one advantage of a constructive proof in general—our proof embodies an algorithm for the construction of the functional whose existence is stated in the theorem. This algorithm can be extracted from the proof, and, as an undeserved bonus, the proof itself demonstrates that the algorithm is correct or, in computer science parlance, "meets its specifications".[1]

So how do we achieve all this? Simply by changing the logic with which we do our mathematics! Instead of using classical logic, we systematically use intuitionistic logic, which was abstracted by Heyting [52] from the practice of Brouwer's intuitionistic mathematics. The remarkable fact is that every proof carried out with intuitionistic logic is fully constructive/algorithmic. (Is this the "secret on the point of being blabbed" that appears in the epigraph to Bishop's book?) Unfortunately, too few mathematicians outside the mathematical logic community are aware of this serendipity and dismiss both intuitionistic logic and constructive mathematics as at best a marginal curiosity. This contrasts sharply with the theoretical computer science community, in which there is considerable knowledge of, and interest in, the computational power of intuitionistic logic.

[1] We do not carry out program-extraction from proofs in our book. For more on this topic see [42, 51, 70].

Reading constructive mathematics demands careful interpretation. A theorem in this book might look like a familiar one from classical analysis, but with more complicated hypotheses and proof. However, the statement of the theorem will be phrased so that the explicit algorithmic interpretation is left to the reader; and the additional hypotheses will be necessary for a constructive proof, which will contain algorithmic information that is excluded from the classical proof by the latter's use of principles outside intuitionistic logic. Consider, for example, the following statement:

> (*) *Let C be an open convex subset of a normed space X, let $\xi \in C$, and let $z \in X$ be bounded away from C. Then the boundary of C intersects the segment $[\xi, z]$ joining ξ and z.*

This is trivial to prove classically; but to find/construct the (necessarily unique) point in which the boundary of C intersects $[\xi, z]$ is a totally different matter. The constructive theorem (Proposition 5.1.5 below) requires us to postulate that the union of C and its metric complement $-C$ (the set of points bounded away from C) be dense in X, and that X itself be a complete normed space. The constructive proof, though elementary, requires some careful geometrical estimation that would be supererogatory in the natural classical proof by contradiction. The benefit of that estimation and of the use of intuitionistic logic is that we could extract from the constructive proof an implementable algorithm for finding the point where the segment crosses the boundary. In turn, this would enable us to produce an algorithm for constructing separating hyperplanes and Hahn–Banach extensions of linear functionals, under appropriate hypotheses.

We could have made the algorithmic interpretation of the constructive version of (*) explicit by stating the proposition in this way:

> *There is a "boundary crossing algorithm" that, applied to the data consisting of* (i) *an open convex set C in a Banach space X such that $C \cup -C$ is dense in X,* (ii) *a point ξ of C, and* (iii) *a point z of $-C$, constructs the point where the boundary of C intersects the segment* $[\xi, z]$.

Even this is not really explicit enough. A full description of the data to which the boundary crossing algorithm applies would require explicit information about the algorithms for such things as these: membership of C; the convergence of Cauchy sequences in X; the computation, for given x in X and $\varepsilon > 0$, of a point y of $C \cup -C$ such that $\|x - y\| < \varepsilon$ (and even the decision between the cases "$y \in C$" and "$y \in -C$"); and so on. Such explicit description of algorithmic hypotheses would become an ever greater burden on writer and reader alike as the book probed deeper and deeper into abstract analysis. It is a matter of sound sense, even sanity, to unburden ourselves from the outset, relying on the reader's native wit in the interpretation of the statements of our constructive lemmas, propositions, and theorems.

We should make it clear that we are not advocating the exclusive use of intuitionistic logic in mathematics. That logic is, we believe, the natural and right one to use when dealing with the constructive content of mathematics. To abandon classical logic in those fields (such as the higher reaches of set theory) where constructivity is of little or no significance makes no sense whatsoever. Nevertheless, it is remarkable how much mathematics actually has what Bishop called “a deep underpinning of constructive truth”.

Christchurch, New Zealand
January 2006

Douglas Bridges
Luminiţa Simona Vîţă

Acknowledgments

It is never easy to apportion thanks properly among the many who have contributed to this book either directly or by their support and encouragement at various stages of our professional lives. We do, however, have special thanks for the following people:

Cris Calude, who was responsible for bringing Vîţă to New Zealand to begin what has proved a very fruitful research partnership with Bridges, and who has been tireless in his encouragement of our work over many years.

Hajime Ishihara, who, in conjunction with one or both of the authors, was largely responsible for much of the work in Chapter 6 and whose influence can be seen in several other places in the book. (Hajime's contributions to constructive functional analysis have been remarkable and deserve to be recognised more widely.) In addition, he has hosted us many times at the Japan Advanced Institute of Science & Technology, each of our visits there being both memorable and highly productive.

Peter Schuster, not only for his contributions to constructive mathematics since he joined our community ten years ago, but also for acting as organiser, fund-raiser, and host on our many research visits to Munich. His considerable efforts in securing a DAAD Gastprofessorship for Bridges in the Mathematisches Institut der Ludwig-Maximilians-Universität (LMU), München, in 2003 provided us with the time and environment in which we could break the back of the writing of this book.

Our early drafts of Chapters 1–5 formed the basis of graduate lectures by Bridges at LMU in 2003. We thank the students in that course for patiently receiving that material and for suggesting corrections and improvements to our presentation of it. We are also grateful to

- the DAAD for supporting Bridges as a Gastprofessor at LMU for that year;
- Otto Forster and Helmut Schwichtenberg, our hosts at LMU in 2003 and on several other occasions;

- the New Zealand Foundation of Research, Science and Technology for Vîţă's postdoctoral fellowship from 2002 to 2005 and for supporting her extended visit to LMU to enable us to work together on the book;
- our departments at Canterbury and Galaţi;
- Josef Berger, Hannes Diener, Maarten Jordans, and Robin Havea, who have kindly assisted us with proofreading various chapters.

We say a warm "thank you" to our friends Imola and Attila Zsigmond, and Helmut and Eva Pellinger, who were wonderful hosts during our time in Munich.

Finally, we want to thank our families for their continuing love and support.

Contents

1

Introduction to Constructive Mathematics

> *My task which I am trying to achieve is, by the power of the written word, to make you hear, to make you feel—it is, before all, to make you see. That, and no more, and it is everything.*
>
> —Joseph Conrad, *The Nigger of the 'Narcissus'*

In this chapter we first sketch the history and philosophy that motivated the early workers in the field of constructive mathematics. We then describe informal intuitionistic logic and discuss a number of elementary classical theorems that do not carry over to the constructive setting. Finally, we introduce an informal constructive theory of sets and functions. All this will prepare us for the presentation of the constructive theory of the real line $\mathbb{R}$ in Chapter 2, and for the more abstract analysis that will be described in later chapters.

1.1 What Is Constructive Mathematics?

Proposition 20 in Book IX of the thirteen volumes of Euclid's *Elements* states that

> *Prime numbers are more than any assigned multitude of prime numbers*

—in current terms, there are infinitely many primes. The modernised version of Euclid's proof is often presented as follows. Suppose that there are only finitely many primes, say $p_1, \ldots, p_n$, and consider the integer

$$p = p_1 \times p_2 \times \cdots \times p_n + 1.$$

Being greater than 2, p has prime factors (it may even be prime itself). Since the numbers p_k are not divisors of p, each prime factor of p is distinct from each p_k. This is absurd, since $\{p_1, \ldots, p_n\}$ is supposed to be the set of *all* primes. From this contradiction we conclude that the set of primes is infinite.

Although at one level there appears to be nothing untoward about this proof, it can be criticised on two counts. First, it uses a totally unnecessary contradiction argument. If you look carefully, you will see that the proof actually embodies an algorithm that, applied to any finite set $\{p_1, \ldots, p_n\}$ of primes, enables you to compute a prime that is distinct from each element of that set. In other words, the use of a contradiction argument in the preceding paragraph has obscured the computational content of Euclid's proof.

The second criticism of the proof is a little more subtle, and deals with the notion of "infinitely many". The proof is based on the negative idea that a set is infinite if and only if it is contradictory that it be finite. But an algorithmic recasting of Euclid's proof, as suggested in the preceding paragraph, shows that the set S of primes is infinite in a more positive, productive sense: namely, if we start with a finite subset F of S, then we can compute an element of S that is distinct from each element of F.

From a traditional standpoint, the distinctions between the contradiction proof of Euclid's theorem and the algorithmic one, and between the negative and positive notions of "infinite", are obscured if not invisible. For example, the two notions of "infinite" are equivalent if we use traditional logic—or *classical logic,* as it is normally called—so the distinction has to be perceived at an aesthetic level rather than a mathematical one. The same applies, more generally, to a proof by contradiction of the existence of an object x with the property $P(x)$. In such a proof one supposes that $P(x)$ is false for all applicable objects x, deduces a contradiction, and then concludes that $P(x)$ must, after all, hold for some x, even though the proof doesn't tell us which x actually has the desired property. Classical logic draws no distinction between the "idealistic existence" demonstrated by such a proof and the "constructive existence" based on an algorithm that constructs x and shows that $P(x)$ holds. In order to reveal such distinctions at a mathematical, rather than an aesthetic, level we shall adopt the radical expedient of changing our logic: *throughout this book, we shall work with intuitionistic logic*, an abstraction of the informal logic used in algorithmic thinking.

How much analysis, as normally presented, is really nonconstructive—that is, essentially dependent on proofs by contradiction or other nonalgorithmic procedures? Consider, for example, the classical *intermediate value theorem*:

> *If* $f : [a, b] \longrightarrow \mathbb{R}$ *is a continuous mapping such that* $f(a) < 0$ *and* $f(b) > 0$, *then there exists* $c \in (a, b)$ *such that* $f(c) = 0$.

(Note that $[a, b]$ and (a, b) respectively denote the closed and open intervals with endpoints a and b. We shall use standard notations, like $(a, b]$ for the half-open interval, without further comment.)

It might be thought that the common elementary proofs of the intermediate value theorem are constructive, enabling one to produce a zero c of the function f. For example, one proof uses interval-halving in the following way. Without loss of generality, take $a = 0$ and $b = 1$. Consider $f(1/2)$: if it is 0, then we take $c = 0$ and stop the process; if $f(1/2) > 0$, then f satisfies the hypotheses of the theorem with $a = 0$ and $b = 1/2$; if $f(1/2) < 0$, then f satisfies the hypotheses with $a = 1/2$ and $b = 1$. In each of the last two cases, we proceed with the interval-halving. This process either stops after a finite number of iterations and produces the required zero of f, or else it goes on ad infinitum to produce a descending sequence of compact intervals whose unique point of intersection is the required zero. Isn't this a fully algorithmic proof?

Suppose we try to implement the algorithm embodied in this proof on a computer that works with 50-bit precision. What happens if we apply it to the cubic function f defined on $[0,1]$ by

$$f(x) = \left(x - \frac{3}{4}\right)\left(x - \frac{1}{2}\right)^2 - 2^{-51}?$$

Here,

$$f(0) = -\frac{3}{16} - 2^{-51} < 0, \quad f(1) = \frac{1}{16} - 2^{-51} > 0,$$

so we are well set to carry out the first step of the interval-halving algorithm. Since our computer's floating-point representation of $f(1/2)$ is 0 (we have the phenomenon of *underflow*, in which the computer sets the small but nonzero number -2^{-51} equal to 0), the algorithm stops by outputting $c = 1/2$ as the place where f has a zero. But in this case the only zero of f in $[0,1]$ lies between $3/4$ and 1, more than one quarter of the entire interval away from the output value $1/2$.

Now, one could object that this example is misleading, in that the problem arises from the level of precision in the computer rather than any intrinsic failing in the algorithm itself. To deal with this point, for each positive integer n let $G(n)$ signify that $2n + 2$ is a sum of two primes. Construct a binary sequence $(a_n)_{n \geqslant 1}$ such that for each n,

> $a_n = 0$ if and only if either $G(k)$ for all $k \leqslant n$ or else there exists $k < n$ such that $\neg G(k)$.

Define $a = \sum_{n=1}^{\infty} a_n 2^{-n}$, and note that $a = 0$ if and only if the *Goldbach conjecture*,

> *Every even integer greater than* 2 *is a sum of two primes,*

holds. Using classical logic, apply the classical interval-halving algorithm to the cubic function f defined on $[0,1]$ by

$$f(x) = \left(x - \frac{3}{4}\right)\left(x - \frac{1}{2}\right)^2 - a.$$

As long as the status of the Goldbach conjecture remains undecided (which it has done since the conjecture first appeared in 1742), no matter what finite precision our computer has, the algorithm will output $1/2$ as a zero of f; but if the Goldbach conjecture is false, then $f(1/2) = -a < 0$ and the first zero of f in $[0,1]$ occurs between $3/4$ and 1. In fact, the classical algorithm will give the correct output if and only if the Goldbach conjecture is true. However, as the reader may verify, a constructive approximate interval-halving argument, such as that expected in the solution of Exercise 11, does not give the possibly false value $1/2$ for a zero of f. It produces a value that approximates the zero of f lying between $3/4$ and 1, as accurately as the precision of the computer permits.

In this example, the classically (but not constructively) defined function taking the parameter a to the smallest root $r(a)$ of f is discontinuous at $a = 0$. The classical algorithm correctly outputs $r(a) = 1/2$ in the case $a = 0$; but if $a > 0$, then, by outputting the value $1/2$, the algorithm has failed to spot that the value of $r(a)$ jumps from $1/2$ to more than $3/4$ as the parameter a increases from 0. There is a general principle that constructive proofs will involve continuity in parameters. Thus we cannot expect to prove constructively that for each a the above cubic function f has a smallest zero.

The problem with the classical interval-halving algorithm is that the finite precision of the computer prevents it from making correct comparisons between two very close, but distinct, real numbers. If we are to develop mathematics in a computational manner, we have to ensure that such comparisons are barred. This barring can be done in at least two ways. One way is to use classical logic and to preclude nonalgorithmic "decisions" (such as whether two given numbers are equal) by developing the mathematics using a standard programming language or a more abstract algorithmic framework like that of recursive function theory. Another way is to change from classical to intuitionistic logic. The advantages of this second way are, first, that nonalgorithmic "decisions" are automatically barred by the logic, and, second, that the resulting mathematics looks like the mathematics we are used to from school and university, without any special logical notation such as is used in, for example, recursive function theory.

In this book we explore mathematics with intuitionistic logic. We work throughout with notions, like that of "infinitely many" discussed earlier, that have positive computational meaning; and we present only algorithmic proofs—ones that show how we can, at least in principle, construct the objects whose existence is asserted in the statement of a theorem. We hope to convince the reader that, contrary to a widely held belief, intuitionistic logic suffices for the development of deep, interesting mathematics and often opens up new vistas that are hidden by classical logic. In other words, we want to justify our belief in the power of positive (constructive) thinking in mathematics.

1.2 A Very Brief History

Although luminaries such as Leopold Kronecker had advocated a constructive approach to mathematics in the nineteenth century, the story of modern constructivism really begins with the publication, in 1907, of the doctoral thesis "On the Foundations of Mathematics" [41], in which the Dutch mathematician L.E.J. Brouwer introduced his *intuitionistic mathematics (INT)* as an alternative to traditional *classical mathematics (CLASS).* According to Brouwer, mathematical objects are free creations of the human mind, independent of both logic and language, and a mathematical object comes into existence precisely when it is constructed. Such a belief naturally leads to a rejection of existence proofs by contradiction, and a consequent scepticism about the meaning of many of the theorems of CLASS. Not surprisingly, Brouwer's views met with at best indifference, and at worst hostility, from the large majority of his peers, for whom the elimination of nonconstructive arguments, with all their apparent power and fruitfulness, was too great a price to pay for a clarification of the meaning of mathematics.

If we adhere to the principle that "existence" should always be interpreted constructively, then we are forced to dispense with the unrestricted use of the logical *law of excluded middle* (or *excluded third*),

$$P \text{ or } (\text{not } P),$$

which we shall abbreviate to *LEM.* Recognising this consequence of his philosophical views, Brouwer went as far as to claim,

> The belief in the universal validity of the principle of the excluded third in mathematics is considered by the intuitionists as a phenomenon of the history of civilization of the same kind as the former belief in the rationality of π, or in the rotation of the firmament about the earth [44].

Subsequently, he introduced into INT some principles that led to results apparently contradicting aspects of classical mathematics. For example, Brouwer was able to prove that any real-valued function on $[0, 1]$ is uniformly continuous. But to regard Brouwer's mathematics as inconsistent with its classical counterpart is a serious oversimplification of the situation, since the two types of mathematics are in many respects incomparable. Nevertheless, there was, and remains, a commonly held belief that too much mathematics has to be given up in order to accommodate Brouwer's ideas. For example, Hilbert expressed his disagreement with Brouwer in words both forceful and memorable:

> Forbidding a mathematician to make use of the principle of excluded middle is like forbidding an astronomer his telescope or a boxer the use of his fists [54].

Despite continuing opposition, intuitionism survived and new constructive approaches to mathematics arose. In 1948–1949 in the former Soviet Union, A.A.

Markov initiated a programme of *recursive constructive mathematics (RUSS)*—mathematics using intuitionistic logic and based on the Church–Markov–Turing thesis that all computable partial functions from the set $\mathbb{N}$ of natural numbers to itself are recursive. This approach led to a number of technical successes [66, 67]. RUSS does not use any of Brouwer's nonlogical intuitionistic principles; indeed, it could not, since it produces results that are false if interpreted directly within INT. For example, in RUSS there exists a continuous real-valued map on $[0,1]$ that is not uniformly continuous; more dramatically, there exists a uniformly continuous map f from $[0,1]$ onto $(0,1]$ that has infimum equal to 0. Once again, one should not overreact to the apparent conflict with classical mathematics: the last of these results should really be interpreted as saying that there exists a recursively uniformly continuous recursive function f from the closed interval $[0,1]$ of the recursive real line onto the recursive interval $(0,1]$ that has infimum equal to 0. Put this way, the result does not conflict with CLASS; indeed, it is a result of CLASS, since the proof within RUSS is actually a proof within CLASS that does not use such nonconstructive logical principles as LEM.

By the mid-1960s, constructive mathematics was, when compared with its classical counterpart, virtually stagnant. The situation changed in 1967 with the publication of Errett Bishop's monograph *Foundations of Constructive Mathematics* [9]. This book and its offspring [12] represent the most far-reaching and systematic presentation of constructive analysis to date. In [9], Bishop revealed, by thoroughgoing constructive means but without resorting to either Brouwer's principles or the formalism of recursive function theory, a vast panorama of constructive mathematics, covering elementary analysis, metric and normed spaces, abstract measure and integration, the spectral theory of selfadjoint operators on a Hilbert space, Haar measure, duality on locally compact groups, and Banach algebras. *Bishop's constructive mathematics* (*BISH*) was founded on a primitive, unspecified notion of "algorithm", or "finite routine", and the Peano properties of natural numbers, and kept strictly to the interpretation of "existence" as "computability". His refusal to pin down the notion of algorithm led to criticism, particularly from philosophers of mathematics and from those committed to the Church–Markov–Turing thesis; but this very imprecision enabled Bishop's work to have a variety of interpretations: his results are valid in CLASS, INT, RUSS, and all reasonable models of computable mathematics, such as the more recent one propounded by Weihrauch [91]. Indeed, from a purely formal viewpoint, each of INT, RUSS, and CLASS can be regarded as BISH plus some additional principles: INT can be regarded as BISH supplemented by Brouwer's continuity principle and fan theorem; RUSS as BISH plus the Church–Markov–Turing thesis; and CLASS as BISH plus the law of excluded middle.

One consequence of this multiplicity of interpretations is that we can often demonstrate that certain propositions P are independent of BISH; that is, neither P nor (not P) can be proved within BISH. For example, since "every mapping from $[0,1]$ into $\mathbb{R}$ is uniformly continuous" is a theorem of INT, and "there exists a continuous map of $[0,1]$ into $\mathbb{R}$ that is not uniformly continuous" is a theorem of

RUSS, and since both INT and RUSS are formally consistent with BISH, within BISH we cannot expect either to prove that every continuous map of $[0,1]$ into $\mathbb{R}$ is uniformly continuous or to construct an example of a real-valued function that is defined, but not uniformly continuous, on $[0,1]$.

Over the years since the publication of Bishop's book, it became clear to a number of researchers that, in essence, BISH is simply *mathematics with intuitionistic logic* together with some appropriate set-theoretic foundation. As we pointed out at the end of the preceding section, working with intuitionistic logic automatically bars noncomputational steps. As long as we keep strictly to intuitionistic logic, having made sure that our set-theoretic principles do not inadvertently imply LEM or some other nonconstructive proposition, the mathematics we develop turns out to be predictive, in the sense that every proof implicitly shows that if we perform certain calculations, we shall achieve certain results. Accordingly, when we speak of "constructive mathematics" or "BISH" in future, we shall mean "mathematics with intuitionistic logic". It therefore behooves us to explain more clearly exactly what intuitionistic logic is.

1.3 Intuitionistic Logic

The meaning doesn't matter if it's only idle chatter of a transcendental kind.

—W.S. Gilbert, *Patience*

Everywhere one seeks to produce meaning, to make the world signify, to render it visible.

—Jean Baudrillard, *Seduction, or the Superficial Abyss*

For Brouwer, mathematics took precedence over logic. In order to describe the logic used by the (intuitionist) mathematician, it was necessary first to analyse the mathematical processes of the mind, from which analysis the logic could be extracted. In 1930, Brouwer's most famous pupil, Arend Heyting (1898–1980), published a set of formal axioms that so clearly characterise the logic used by the intuitionist that they have become universally known as the *axioms for intuitionistic logic* [52]. These axioms capture the so-called *BHK interpretation* of the connectives

$$\vee \text{ (or)},\ \wedge \text{ (and)},\ \Longrightarrow \text{ (implies)},\ \neg \text{ (not)}$$

and quantifiers

$$\exists \text{ (there exists)}, \quad \forall \text{ (for all/each)},$$

which we now outline.

- $P \vee Q$: either we have a proof of P or else we have a proof of Q.
- $P \wedge Q$: we have both a proof of P and a proof of Q.
- $P \Longrightarrow Q$: by means of an algorithm we can convert any proof of P into a proof of Q.
- $\neg P$: assuming P, we can derive a contradiction (such as $0 = 1$); equivalently, we can prove $(P \Longrightarrow (0 = 1))$.
- $\exists x\, P(x)$: we have (i) an algorithm that computes a certain object x, and (ii) an algorithm that, using the information supplied by the application of algorithm (i), demonstrates that $P(x)$ holds.
- $\forall x \in A\, P(x)$: we have an algorithm that, applied to an object x and a proof that $x \in A$, demonstrates that $P(x)$ holds.

Note that in the interpretation of the statement $\forall x \in A\, P(x)$, the proof of $P(x)$ will normally use *both* the data describing the object x *and* the information supplied by a proof that x belongs to the set A. This is an important point, since upon it hinges a key argument against the use of the axiom of choice in constructive mathematics. We shall return to this matter later.

A property $P(x)$ is said to be *decidable* if for each x to which it might be applicable we have

$$P(x) \vee \neg P(x),$$

where the disjunction and negation are given their BHK interpretations. Even for a decidable property $P(n)$ of natural numbers n the property

$$\forall n\, P(n) \vee \neg \forall n\, P(n),$$

and hence a fortiori LEM, will not hold in general. As a result, many classical results cannot be proved constructively, since they would imply LEM or perhaps some other manifestly nonconstructive principle.

To illustrate this point, consider the following simple statement, the *limited principle of omniscience (LPO)*:

$$\forall \mathbf{a} \in \{0,1\}^{\mathbb{N}^+} (\mathbf{a} = \mathbf{0} \vee \mathbf{a} \neq \mathbf{0}),$$

where $\mathbf{a} = (a_1, a_2, \ldots)$, $\mathbb{N}^+ = \{1, 2, \ldots\}$ is the set of positive integers, $\{0,1\}^{\mathbb{N}^+}$ is the set of all binary sequences, and

$$\begin{aligned} \mathbf{a} = \mathbf{0} &\Longleftrightarrow \forall n\, (a_n = 0), \\ \mathbf{a} \neq \mathbf{0} &\Longleftrightarrow \exists n\, (a_n = 1). \end{aligned}$$

In words, LPO states that for each binary sequence $(a_n)_{n \geqslant 1}$, either $a_n = 0$ for all n or else there exists n such that $a_n = 1$. Of course, this is a triviality from the viewpoint of classical logic. But its BHK interpretation is not so simple: it says that there is an algorithm that, applied to any binary sequence **a**, either verifies that all the terms of the sequence are 0 or else computes the index of a term equal to 1. Anyone familiar with computers ought to be highly sceptical about such an algorithm, since in the case $\mathbf{a} = \mathbf{0}$ it would normally need to test each of the infinitely many terms a_n in order to come up with the correct decision.

For such reasons we feel justified in not accepting LPO, or any classical proposition that constructively implies LPO, as a valid principle of constructive mathematics. But we have another reason for not doing so: it can be shown that there are models of *Heyting arithmetic*—Peano arithmetic with intuitionistic logic—in which LPO is false; so LPO cannot be derived in Heyting arithmetic (see [34, 48]). Since LPO is a special case of the law of excluded middle, we are led, in turn, to renounce the latter when working constructively. Similar informal analyses lead us to exclude both the classical rule

$$\neg\neg P \Longrightarrow P,$$

which forms the basis of proof by contradiction, and the following *lesser limited principle of omniscience (LLPO)*, which is easily seen to be a consequence of LPO:

> *For each binary sequence* **a** *with at most one term equal to* 1 *(in the sense that* $a_m a_n = 0$ *for all distinct* m *and* n*), either* $a_{2n} = 0$ *for all* n *or else* $a_{2n+1} = 0$ *for all* n.

The exclusion of such principles from constructive mathematics has serious consequences for mathematical practice. For example, we cannot hope to prove constructively the simple statement

$$\forall x \in \mathbb{R}\ (x = 0 \vee x \neq 0)\,, \tag{1.1}$$

where $\mathbb{R}$ denotes the set of real numbers and $x \neq 0$ means that we can compute a rational number strictly between 0 and x (which, as we shall see when we deal with the real numbers more formally in Chapter 2, is not the same, constructively, as proving that $\neg\,(x = 0)$). To prove this, consider any binary sequence **a**, and use it to define the binary expansion of a real number

$$x = \sum_{n=1}^{\infty} a_n 2^{-n}.$$

If $x = 0$, then $\mathbf{a} = \mathbf{0}$. If $x \neq 0$, we can compute a positive integer N such that

$$x > 2^{-N} = \sum_{n=N+1}^{\infty} 2^{-n};$$

it is then clear that, by testing the terms $a_1, \ldots, a_N$, we can find $n \leqslant N$ such that $a_n = 1$. Thus statement (1.1) about real numbers implies LPO and is therefore essentially nonconstructive.

If the binary sequence $\mathbf{a}$ has at most one term equal to 1, then we can use the real number

$$\sum_{n=1}^{\infty} (-1)^n \, a_n 2^{-n}$$

to show that the statement

$$\forall x \in \mathbb{R} \; (x \geqslant 0 \vee x \leqslant 0)$$

implies LLPO.

The following elementary classical statements also turn out to be nonconstructive.

- ▷ Each real number x is either rational or irrational (that is, $x \neq r$ for each rational number r). To see this, consider

$$x = \sum_{n=1}^{\infty} \frac{1 - a_n}{n!},$$

  where $\mathbf{a}$ is any increasing binary sequence (that is a binary sequence such that $a_n \leqslant a_{n+1}$ for each n).

- ▷ Each real number x has a binary expansion. Note that the standard interval-halving argument for "constructing" binary expansions does not work, since we cannot necessarily decide, for a given number x between 0 and 1, whether $x \geqslant 1/2$ or $x \leqslant 1/2$. In fact, the existence of binary expansions is equivalent to LLPO.

- ▷ The intermediate value theorem, which is equivalent to LLPO.

- ▷ For all $x, y \in \mathbb{R}$, if $xy = 0$, then either $x = 0$ or $y = 0$. The constructive failure of this proposition clearly has implications for the theory of integral domains.

We emphasise here that classically valid statements like "each real number is either rational or irrational" that imply omniscience principles are *not false* in constructive mathematics; they cannot be, since BISH is consistent with CLASS.

One principle whose constructive status is controversial is *Markov's principle (MP):*

$$\forall \mathbf{a} \in \{\mathbf{0}, \mathbf{1}\}^{\mathbb{N}^+} \; (\neg (\mathbf{a} = \mathbf{0}) \Longrightarrow \mathbf{a} \neq \mathbf{0});$$

in words, for any binary sequence $\mathbf{a}$, if it is impossible for all the terms to equal 0, then there exists a term equal to 1. In order to accept this as a principle of constructive mathematics, you have to be convinced that the information conveyed by the antecedent $\neg (\mathbf{a} = \mathbf{0})$ is sufficient to enable us to compute an index n with $a_n = 1$. The argument in favour of MP says that we can carry out this computation by searching systematically through the terms a_n, since the hypothesis $\neg (\mathbf{a} = \mathbf{0})$

guarantees that we shall eventually stumble across a term equal to 1. The counter-argument is that the antecedent provides us with no prior bound for such a search—it does not tell us how many terms we need to test before we arrive at one equal to 1—so the search might go on longer than the remaining life of the universe before it produced the desired result. Moreover, Markov's principle, like LPO, cannot be proved within Heyting arithmetic. For these reasons, we shall follow the normal practice of excluding MP from the working principles of constructive mathematics. As a consequence we exclude the even stronger logical principle

$$(\forall x \in A\,(P(x) \vee \neg P(x)) \wedge \neg\forall x \in A\,\neg P(x)) \Longrightarrow \exists x \in A\,P(x), \tag{1.2}$$

where A is a well-defined set (the exact meaning of "well-defined set" will become clear in the next section). In fact, even if we were to accept Markov's principle on the grounds that an unbounded search through the natural numbers that cannot fail to terminate must eventually do so, we would balk at accepting (1.2), since for a general set A there will be no natural order allowing us to search systematically in the way we can with $\mathbb{N}$.

An example of the type dealt with earlier, in which a classically valid proposition P is shown constructively to entail an essentially nonconstructive principle like LEM, LPO, LLPO, or even MP, is called a *Brouwerian counterexample to P* (even though it is not a counterexample in the true sense of the word; it is merely an indication that P does not admit of constructive proof). There is another expression that we may use in this context. For example, we refer to the number $x = \sum_{n=1}^{\infty} a_n 2^{-n}$ that we constructed from a given binary sequence $(a_n)_{n\geqslant 1}$ and then used to show that (1.1) implies LPO as a *Brouwerian example of a real number x for which we cannot decide whether* $x = 0$ or $x \neq 0$.

1.4 Informal Set Theory

> Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.
>
> —L. Kronecker [90]

> The primary concern of mathematics is number, and this means the positive integers. We feel about number the way Kant felt about space. The positive integers and their arithmetic are presupposed by the very nature of our intelligence and, we are tempted to believe, by the very nature of intelligence in general. The development of the positive integers from the primitive concept of the unit, the concept of adjoining a unit, and the process of mathematical induction carries complete conviction. In the words of Kronecker, the positive integers were created by God.
>
> —Errett Bishop [9]

Building on the set of positive integers and using intuitionistic logic, we follow Bishop's approach to developing constructive mathematics at higher and higher levels of abstraction. To do this, we need to clarify notions such as "set" and "function".

For us, a *set* X (other than our basic, primary set $\mathbb{N}^+$ of positive integers) is given by two pieces of data:

- ▷ a property that enables members of X to be constructed using objects that have already been constructed (note this last phrase, which rules out the possibility of impredicative definitions and therefore of Russell-type paradoxes), and
- ▷ an equivalence relation $=_X$ of *equality* between members of X.

We write $x \in A$ to signify that x is an element of the set A, and $x \notin A$ instead of $\neg (x \in A)$.

The use of equivalence relations rather than intensional equality—that is, identity of description—is common, but often goes unnoticed, in classical mathematics. For example, we call the rational numbers $1/2$ and $3/6$ equal, even though, strictly speaking, they are equivalent and not intensionally identical.

A *subset* S of a set X consists of a collection of elements drawn from X, together with the equality relation induced on S by the given equality on X; that is, for elements x, y of S, we define

$$x =_S y \Longleftrightarrow x =_X y.$$

We write $S \subset T$ to signify that S is a subset of T. If $P(x)$ is a property applicable to certain elements x of a set A, then we denote by

$$\{x : x \in A \wedge P(x)\}$$

or

$$\{x \in A : P(x)\}$$

the subset of A consisting of those elements x of A with the property $P(x)$.

The *logical complement* of a subset S of X is

$$\neg S = \{x \in X : x \notin S\}.$$

A particular example of this is the *empty subset* of X, defined by

$$\varnothing_X = \neg X.$$

We say that a subset S of X is *inhabited* if

$$\exists x \, (x \in S).$$

We then write $S \neq \varnothing_X$. Note that in order to show that S is inhabited, we cannot just prove that it is impossible for S to be empty; we must actually construct an element of S; see Exercise 2.

Two sets X, Y are said to be *equal sets* if each is a subset of the other; in other words, if the sets have the same elements and the same equality relation.

We need to be careful when constructing new sets from old. Since an equality is part of the data for a set, it does not make sense to talk of the union $S \cup T$ of two sets unless we can put together not only the sets as collections of objects but also, in some way, their given equality relations. In practice, this means that in order to construct their *union* $S \cup T$, the sets S and T must be given as subsets of some set X. We then define

$$S \cup T = \{x \in X : x \in S \vee x \in T\},$$

where the equality on $S \cup T$ is that induced by X. Likewise, the *intersection* of S and T is defined only when S and T are subsets of some set X, and is then the subset

$$S \cap T = \{x \in X : x \in S \wedge x \in T\}$$

of X.

The *(Cartesian) product* of two sets X, Y is the set $X \times Y$ consisting of all ordered pairs (x, y) with $x \in X$ and $y \in Y$, together with the equality given by

$$((x, y) =_{X \times Y} (x', y')) \Longleftrightarrow (x =_X x' \wedge y =_Y y')\,.$$

In many situations—even, as we shall see, on the real line—we frequently need a set X to be equipped with an *inequality* relation $\neq_X$ describing what it means for two elements of X to be *unequal*, or *distinct.* Such a relation must satisfy the following two properties:

$$\begin{aligned} x \neq_X y &\Longrightarrow \neg\,(x =_X y)\,, \\ x \neq_X y &\Longrightarrow \quad y \neq_X x. \end{aligned}$$

If, in addition,

$$\neg\,(x \neq_X y) \Longrightarrow x =_X y,$$

we say that the inequality is *tight.* A set X with an inequality is *discrete* if, for any two elements x and y of X, either $x =_X y$ or $x \neq_X y$; we then also describe the inequality itself as discrete.

One inequality relation, the *denial inequality,* is defined by setting $x \neq_X y$ if and only if $\neg\,(x =_X y)$. This inequality is normally too weak for practical purposes. For example, in the absence of Markov's principle, on the real line $\mathbb{R}$ the property $\neg\,(x =_{\mathbb{R}} 0)$ is weaker than $|x| > 0$ (Exercise 3); for that reason we define the standard inequality on $\mathbb{R}$ to be not the denial inequality but the one given by

$$x \neq_{\mathbb{R}} y \Longleftrightarrow |x - y| > 0.$$

From now on, when the meaning is clear from the context, we write

$$=,\ \varnothing,\ \neq,\ \ldots$$

rather than

$$=_X, \ \varnothing_X, \ \neq_X, \ \ldots .$$

A subset S of a set X with an inequality has a *complement*, defined by

$$\sim S = \{x \in X : \forall s \in S \, (x \neq s)\} .$$

Then $\sim S \subset \neg S$; but unless the inequality on X is the denial inequality, the reverse inclusion will not hold.

A subset S of a set X is said to be *detachable* (from, or in, X) if

$$\forall x \in X \, (x \in S \vee x \notin S) .$$

Since statement (1.1) implies LPO, not even the singleton subset $\{0\}$ is detachable from $\mathbb{R}$. However, $\{0\}$ is detachable in the set $\mathbb{Q}$ of rational numbers.

When X comes with an inequality relation, we define the inequality on any subset S of X to be the one induced by that on X:

$$x \neq_S y \Longleftrightarrow x \neq_X y.$$

If also Y has an inequality relation, we define the inequality on the Cartesian product $X \times Y$ by

$$((x, y) \neq_{X \times Y} (x', y')) \Longleftrightarrow (x \neq_X x' \vee y \neq_Y y') .$$

It should be no surprise that we require functions to be given by algorithms and to respect equality. Thus a *function* f from a set X to a set Y—also called a *map* or *mapping* of X into Y, and written $f : X \longrightarrow Y$—is an algorithm that, applied to any element x of X, produces an element $f(x)$ of Y such that f is *extensional*:

$$\forall x \in X \, \forall x' \in X \, (x =_X x' \Longrightarrow f(x) =_Y f(x')) .$$

The element $f(x)$ is called the *value of f at x* or the *image of x under f*. If X and Y have inequality relations, then we may require f to be *strongly extensional*:

$$\forall x \in X \, \forall x' \in X \, (f(x) \neq_Y f(x') \Longrightarrow x \neq_X x') .$$

Note that the statement "all functions from $\mathbb{R}$ to $\mathbb{R}$ are strongly extensional" is equivalent to Markov's principle.

Let f, g be two real-valued functions on a set X with an inequality. We say that an element x_0 of X is the *strongly unique element of X such that* $f(x) = g(x)$ if $f(x_0) = g(x_0)$ and $f(x) \neq g(x)$ whenever $x \in X$ and $x \neq x_0$.

Strong uniqueness will resurface in Chapter 4 in connection with best approximations.

A *partial function* $f : X \longrightarrow Y$ is a function from a subset of X into Y. The subset

$$\{x \in X : f(x) \text{ is defined}\}$$

is called the *domain* of f, denoted by $\operatorname{dom}(f)$; and the set

$$\{y \in Y : \exists x \in X\, (y =_Y f(x))\}$$

the *range* of f, denoted by $\operatorname{ran}(f)$. The *image of a subset A of X under f* is the set

$$f(A) = \{f(x) : x \in A\}.$$

The *inverse image* of a subset B of Y under f is the set

$$f^{-1}(B) = \{x \in \operatorname{dom}(f) : f(x) \in B\}.$$

A partial function $f : X \longrightarrow Y$ is said to be a *total partial function* on X if $\operatorname{dom}(f) = X$.

When the expression describing $f(x)$ is given explicitly and the domain of the partial function f is clearly understood, we may denote the function by $x \rightsquigarrow f(x)$. For example, the partial function from $\mathbb{R}$ to itself whose value is defined, for each real number x such that $x \neq_{\mathbb{R}} 0$, to be $1/x$ may be written $x \rightsquigarrow 1/x$.

An important type of total partial function is defined as follows. Let X and I be sets. A *family of elements of X* with *index set I* (or *indexed by I*) is a mapping $i \rightsquigarrow x_i$ of I into X; we commonly denote this family by $(x_i)_{i \in I}$. In particular, if I is $\mathbb{N}^+$, the family is called a *sequence in X* and is usually written $(x_n)_{n \geqslant 1}$. More general sequences of the form $(x_n)_{n \geqslant N}$ have the obvious analogous meaning when N is an integer.

Let f and g be mappings from subsets of a set X into a set Y, where Y is equipped with a binary operation $\diamond$. We introduce the corresponding *pointwise operation* $\diamond$ on f and g by setting

$$(f \diamond g)(x) = f(x) \diamond g(x)$$

whenever $f(x)$ and $g(x)$ are both defined. Thus, taking $Y = \mathbb{R}$, we see that the (pointwise) sum of f and g is given by

$$(f + g)(x) = f(x) + g(x)$$

if $f(x)$ and $g(x)$ are both defined; and that the (pointwise) quotient of f and g is given by

$$(f / g)(x) = f(x) / g(x)$$

if $f(x)$ and $g(x)$ are defined and $g(x) \neq 0$. If $X = \mathbb{N}^+$, so that $f = (x_n)_{n \geqslant 1}$ and $g = (y_n)_{n \geqslant 1}$ are sequences, then we also speak of *termwise operations*; for example, the *termwise product* of f and g is the sequence $(x_n y_n)_{n \geqslant 1}$.

Pointwise operations extend in the obvious ways to finitely many functions. In the case of a sequence $(f_n)_{n \geqslant 1}$ of functions with values in a normed space (see

Chapter 2), once we have introduced the notion of a series in a normed space we shall interpret $\sum_{n=1}^{\infty} f_n$ in the obvious pointwise way.

A partial function $f : X \longrightarrow Y$ can be identified with its *graph*,

$$\mathcal{G}(f) = \{(x, y) \in X \times Y : x \in \operatorname{dom}(f), \, y \in \operatorname{ran}(f), \, y = f(x)\},$$

a subset of the Cartesian product $X \times Y$. We define two partial functions $f : X \longrightarrow Y$ and $g : X \longrightarrow Y$ to be *equal* if their graphs are equal as sets. Thus f, g are equal partial functions if and only if their domains are equal subsets of X and $f(x) = g(x)$ for each $x \in \operatorname{dom}(f)$.

We say that a partial function $f : X \longrightarrow Y$ is

- ▶ *one-one* if $f(x) = f(x')$ entails $x = x'$;
- ▶ *injective* if X and Y are equipped with inequality relations, and $x \neq x'$ entails $f(x) \neq f(x')$.

If f is injective and the inequality on X is tight, then f is one-one: for in that case, if $f(x) = f(x')$, then $\neg (x \neq x')$ and so, by tightness, $x = x'$.

The *composition*, or *composite*, of partial functions $f : A \longrightarrow B$ and $g : B \longrightarrow C$ is the partial function $g \circ f : A \longrightarrow C$ (sometimes written gf) defined by $g \circ f(x) = g(f(x))$ wherever the right side exists.

A partial mapping $f : X \longrightarrow Y$ maps its domain *onto* Y if

$$\forall y \in Y \exists x \in X \, (y = f(x)) \, .$$

On the other hand, we say that $f : X \longrightarrow Y$ is an *epimorphism* if there exists a mapping $g : Y \longrightarrow X$ such that

$$\forall y \in Y \, (f \, (g(y)) = y) \, .$$

A one-one partial function $f : X \longrightarrow Y$ has a one-one *inverse*

$$f^{-1} : \operatorname{ran}(f) \longrightarrow \operatorname{dom}(f)$$

defined by $f^{-1}(f(x)) = x$. If f is injective, then its inverse is strongly extensional. A *bijection* between X and Y is a one-one mapping from X onto Y.

The subtle distinction between mappings onto and epimorphisms is closely linked to the constructive status of the axiom of choice, which we shall discuss shortly. First, though, we introduce some notions of cardinality.

Let S be a subset of a set X with an inequality relation. We say that S is

- ▷ *finitely enumerable* if there exist a natural number N and a mapping of the set

$$\{1, 2, \ldots, N\} = \{n \in \mathbb{N}^+ : n \leqslant N\}$$

onto S;

▷ *finite* if there exist a natural number N and an injective map of $\{1, 2, \ldots, N\}$ onto S.

When we speak of *finitely many* objects, we mean that those objects constitute an inhabited, finitely enumerable, but not necessarily finite, set.

Note that the case $N = 0$ of the definition of "finitely enumerable" shows that the empty subset of X is finitely enumerable. Clearly, finite implies finitely enumerable; the converse does not hold constructively.

We say that S is *countable* if there exist a detachable subset D of $\mathbb{N}^+$ and a mapping ϕ of D onto S. Every finitely enumerable subset—in particular, the empty subset—of X is countable. If $D = \mathbb{N}^+$, we call ϕ an *enumeration* of S, in which case we may denote S by $\{\phi(1), \phi(2), \ldots\}$. If also ϕ is one-one, we say that S is *countably infinite*. A set is countably infinite if and only if it is the range of a one-one mapping whose domain is a countably infinite, detachable subset of $\mathbb{N}^+$.

We now consider the *axiom of choice (AC):*

> *If X, Y are inhabited sets, S is a subset of $X \times Y$, and for each $x \in X$ there exists $y \in Y$ such that $(x, y) \in S$, then there exists a* **choice function** *$f : X \longrightarrow Y$ such that $(x, f(x)) \in S$ for each $x \in X$.*

Under the BHK interpretation, the hypothesis

$$\forall x \in X \; \exists y \in Y \;\; ((x, y) \in S)$$

of AC means that we have an algorithm that, applied to each element x of X and the data showing that x belongs to X, constructs an element y of Y and demonstrates that $(x, y) \in S$. This much is clear. However, there is no guarantee that the algorithm will respect the equality relation on X—in other words, that if $x =_X x'$, and the algorithm constructs y, y' in Y such that $(x, y) \in S$ and $(x', y') \in S$, then $y =_Y y'$. Indeed, we should expect that the computation of y might use data that are associated with properties intrinsic to x that do not apply intrinsically to x'.

For example, anticipating our development of the real number set $\mathbb{R}$, consider the case in which X is $\mathbb{R}$ and Y is $\mathbb{N}^+$. A real number is (defined as) a certain set of rational approximations. However, two equal real numbers x, x' can have different defining sets of rational approximations. In that case, the algorithm that computes a positive integer n such that $(x, n) \in S$ may, and in general will, compute a different positive integer n' such that $(x', n') \in S$. These considerations throw real doubt over the possibility that there is a choice function implementing the algorithm.

In fact, an argument of Diaconescu [46] and Goodman & Myhill [50], but prefigured by Bishop (see Problem 2 on page 58 of [9]), shows that AC cannot be allowed as a principle of constructive mathematics.

Theorem 1.4.1. *The axiom of choice implies the law of excluded middle.*

Proof. Let P be any constructively meaningful statement, and define the set X to consist of the two elements 0 and 1, together with the equality relation such that

$$(0 =_X 1) \Longleftrightarrow P.$$

Let Y be the set $\{0,1\}$ with the standard equality, and let S be the subset $\{(0,0),(1,1)\}$ of the Cartesian product $X \times Y$, taken with the standard equality. Suppose there exists a function $f : X \longrightarrow Y$ such that $(x, f(x)) \in S$ for all $x \in X$. There are three cases to consider: (i) $f(0) = 1$, (ii) $f(1) = 0$, and (iii) both $f(0) = 0$ and $f(1) = 1$. In case (i) we have $(0,1) = (0, f(0)) \in S$, so either $(0,1) =_{X\times Y} (0,0)$ or $(0,1) =_{X\times Y} (1,1)$. If the first of these two alternatives holds, then, by definition of the equality on $X \times Y$, we have $1 =_Y 0$, which is absurd. Hence, in fact, $(0,1) =_{X\times Y} (1,1)$. Thus, again by definition of the equality on $X \times Y$, we have $0 =_X 1$ and therefore P holds. Case (ii) similarly leads to the conclusion that P holds. Finally, in case (iii) we have $\neg\,(f(0) =_Y f(1))$; therefore, since f is a function, $\neg\,(0 =_X 1)$ and so $\neg P$ holds. Thus we have derived $P \vee \neg P$ from AC. □

The axiom of choice will hold constructively if the set X is one for which no computation is necessary to demonstrate that an element belongs to it; Bishop calls such sets *basic sets.* Following the practice of most constructive mathematicians, we consider $\mathbb{N}^+$, the set

$$\mathbb{N} = \{0,1,2,3,\ldots\}$$

of natural numbers, and the set

$$\mathbb{Z} = \{0,\pm1,\pm2,\ldots\}$$

of all integers to be basic sets. This practice is reflected in our acceptance of the *principle of countable choice:*

> *If Y is an inhabited set, S is a subset of $\mathbb{N}^+ \times Y$, and for each positive integer n there exists $y \in Y$ such that $(n, y) \in S$, then there is a function $f : \mathbb{N}^+ \longrightarrow Y$ such that $(n, f(n)) \in S$ for each $n \in \mathbb{N}^+$.*

In fact, many constructive proofs use the stronger *principle of dependent choice:*

> *If X is a set, $a \in X$, S is a subset of $X \times X$, and for each $x \in X$ there exists $y \in X$ such that $(x, y) \in S$, then there exists a sequence $(x_n)_{n\geqslant1}$ in X such that $x_1 = a$ and $(x_n, x_{n+1}) \in S$ for each $n \in \mathbb{N}^+$.*

Another contentious matter in constructive mathematics is the status and role of the *power set* $\mathcal{P}(X)$ of a given set X: that is, the collection of all subsets of X, with equality of subsets as defined earlier. The main objection to admitting $\mathcal{P}(X)$ into the constructive fold is that we thereby allow impredicativity, since there is then nothing to stop us constructing subsets of X whose defining characteristics are self-referential. On the other hand, nobody has yet shown that adding the power set axiom

$$\forall X \exists Y \forall S\,(S \subset X \Longleftrightarrow S \in Y)$$

to constructive mathematics enables us to prove LEM or some other incontestably nonconstructive principle.

There are ways of avoiding the power set. It often suffices to work with the set Y^X of all mappings from X into a set Y. Since we have a clear idea of what mappings from X into Y are (something we do not have for subsets of X), the set Y^X seems relatively innocent. Note that classically the set $\{0,1\}^X$ can be identified with the power set of X, since it comprises the characteristic functions of subsets of X. This identification is not possible constructively, since characteristic functions exist only for those subsets of X that are detachable.

Another way to avoid the full generality of the power set is to work with a well-defined but smaller set of subsets of X. For example, the set of compact subsets of a metric space X is well defined (it is actually a metric space itself), and is often all we need for many parts of analysis.

Let $\mathcal{S}(X)$ be a well-defined set of subsets of X, with two elements taken as equal if and only if they are equal sets in the usual sense, and let I be some set. Then we can speak sensibly about a family $(S_i)_{i \in I}$ of elements of $\mathcal{S}(X)$. We can also define the *union* and *intersection* of such a family to be, respectively, the subsets

$$\bigcup_{i \in I} S_i = \{x \in X : \exists i \in I\,(x \in S_i)\}$$

and

$$\bigcap_{i \in I} S_i = \{x \in X : \forall i \in I\,(x \in S_i)\}$$

of X. If I is the set of positive integers, we denote the above union and intersection by $\bigcup_{n \geqslant 1} S_n$ and $\bigcap_{n \geqslant 1} S_n$, respectively.

The *Cartesian product* $\prod_{i \in I} S_i$ is the subset of X^I consisting of those functions f such that $f(i) \in S_i$ for each $i \in I$; if X comes with an inequality relation, then the corresponding inequality on $\prod_{i \in I} S_i$ is defined by

$$f \neq g \Longleftrightarrow \exists i \in I\,(f(i) \neq g(i))\,.$$

If I is the set of positive integers, then $(S_i)_{i \in I}$ is a sequence of sets, and the elements of $\prod_{i \in I} S_i$ are also sequences; in this case, the sets S_i can be arbitrary and need not be subsets of a previously defined set X. If $I = \{1, 2, \ldots, n\}$, then we denote the product $\prod_{i \in I} S_i$ by $S_1 \times S_2 \times \cdots \times S_n$, and refer to its elements $(x_1, x_2, \ldots, x_n)$ as *ordered n-tuples,* or, in the case $n = 2$, *ordered pairs* (something we used informally in our earlier definition of the Cartesian product of two sets).

Other set-theoretic notions will be introduced later as they arise. It is now time to turn our attention away from foundational matters to analysis proper.

Exercises

Although we shall not formally define the real number line $\mathbb{R}$ until Chapter 2, in these exercises we assume elementary properties of real numbers where necessary.

1. Justify informally Brouwer's observation that $\neg\neg\neg P$ implies $\neg P$. Using this, show that the proposition $(\neg\neg P \Longrightarrow P)$ is equivalent to the law of excluded middle.

2. Prove that if the statement
$$\neg (S = \varnothing) \Longrightarrow S \neq \varnothing$$
applies to all subsets S of $\{0\}$, then the law of excluded middle holds.

3. Prove that
$$\forall x \in \mathbb{R}\, (\neg (x = 0) \Longrightarrow x \neq 0)$$
is equivalent to Markov's principle.

4. Fill in the details of the proof that the statement
$$\forall x \in \mathbb{R}\ (x \geqslant 0 \vee x \leqslant 0)$$
implies LLPO.

5. Give a Brouwerian counterexample to the proposition that every real number x is either rational or irrational (where "x is irrational" means that $x \neq r$ for each rational number r).

6. Prove that every real number has a binary expansion if and only if LLPO holds.

7. Give a Brouwerian counterexample to the statement that if r_1, r_2, and r_3 are real roots of a quadratic polynomial $x^2 + ax + b$ with $a, b \in \mathbb{R}$, then there exist distinct i, j with $r_i = r_j$.

8. Prove that the statement "for all real numbers x, y that have binary expansions, the sum $x + y$ has a binary expansion" is equivalent to LLPO.

9. Prove that the statement

$$\forall x, y \in \mathbb{R}\ (xy = 0 \Longrightarrow (x = 0 \vee y = 0))$$

is equivalent to LLPO.

10. Prove that the intermediate value theorem is equivalent to LLPO.

11. Let $f : [a, b] \longrightarrow \mathbb{R}$ be *sequentially continuous* in the following sense: for each sequence $(x_n)_{n\geqslant 1}$ in $[a, b]$ that converges to a limit x, the sequence $(f(x_n))_{n\geqslant 1}$ converges to $f(x)$. Suppose also that $f(a)f(b) < 0$, and that f is *locally nonzero* in the sense that for each $x \in [a, b]$ and each $r > 0$ there exists $y \in [a, b]$ with $|x - y| < r$ and $f(y) \neq 0$. Prove that there exists $c \in (a, b)$ such that $f(c) = 0$. (This version of the intermediate value theorem suffices for virtually all constructive purposes.)

12. Prove that the statement "all functions from $\mathbb{R}$ to $\mathbb{R}$ are strongly extensional" is equivalent to Markov's principle.

13. Discuss the statement "every mapping from a set X onto a set Y is an epimorphism".

14. Give a Brouwerian counterexample to the statement "every subset of a finite set is finitely enumerable".

15. Prove that an inhabited set is countably infinite if and only if it is the range of a function with domain $\mathbb{N}$.

16. Prove that a subset S of $\mathbb{N}$ is countable if and only if there exists a sequence $(S_n)_{n\geqslant 1}$ of finite subsets of $\mathbb{N}$ such that $S_1 \subset S_2 \subset S_3 \subset \cdots$.

17. Prove that the statement "every inhabited subset of $\mathbb{N}$ is countable" implies the *principle of finite possibility (PFP):* to each binary sequence $(a_n)_{n\geqslant 1}$ there corresponds a binary sequence $(b_n)_{n\geqslant 1}$ such that $a_n = 0$ for each n if and only if there exists N such that $b_N = 1$.

18. Prove that the intersection of two countable sets is countable.

19. Prove that the principle of dependent choice implies the principle of countable choice. Prove also that the principle of dependent choice can be derived from the following *principle of internal choice:* if X is an inhabited set, S is a subset of $X \times X$, and for each $x \in X$ there exists $y \in X$ such that $(x, y) \in S$, then there exists a choice function $f : X \longrightarrow X$ such that $(x, f(x)) \in S$ for each $x \in X$.

20. By a *filter* on a set X we mean a set $\mathfrak{F}$ of inhabited subsets of X with the following properties:

    - $X \in \mathfrak{F}$.
    - If S and T belong to $\mathfrak{F}$, then $S \cap T \in \mathfrak{F}$.
    - If $S \in \mathfrak{F}$ and $S \subset T \subset X$, then $T \in \mathfrak{F}$.

    A filter $\mathfrak{U}$ is called an *ultrafilter* if for all $S \subset X$, either $S \in \mathfrak{U}$ or $\neg S \in \mathfrak{U}$. The classical *ultrafilter principle* states that every filter is contained in an ultrafilter. Prove that this principle implies the *weak limited principle of omniscience (WLPO):*

$$\forall \mathbf{a} \in 2^{\mathbb{N}} \left(\forall n \left(a_n = 0 \right) \vee \neg \forall n \left(a_n = 0 \right) \right).$$

Notes

The use of the Goldbach conjecture in the example in Section 1.1 is in no sense a constraint on the argument; if the Goldbach conjecture were solved tomorrow, we could easily replace it by any one of many open problems of number theory. Indeed, until Wiles proved the Fermat conjecture in 1994, it was common to use that, rather than the Goldbach conjecture, in such Brouwerian examples of nonconstructivity.

The reader interested in the history of foundations of mathematics should consult [44] and [84] for more information on the "Grundlagenstreit" that eventuated between Brouwer and Hilbert in the 1920s. For more on Hilbert's programme for the secure foundation of mathematics see [93].

The designation "BHK interpretation" comes from **B**rouwer, **H**eyting, and **K**olmogorov. The BHK interpretation of implication, while more natural than the classical one of *material implication,* in which $(P \Longrightarrow Q)$ is equivalent to $(\neg P \vee Q)$, has not completely satisfied all researchers using constructive logic. Shortly before he died, Bishop communicated to Bridges his dissatisfaction with the standard constructive interpretation of implication. Unfortunately, he left nothing more than very rudimentary sketches of his ideas for its improvement. (Note, however, Bishop's paper [10].) For a deeper analysis of the constructive interpretations of the connectives and quantifiers, we refer the reader to [48].

The classical invalidity of the recursive interpretation of LPO is not a matter of logic: it can be demonstrated, even with classical logic, that a recursive version of LPO would lead to a proof of the decidability of the halting problem, which is known to be impossible; see [34], pages 52–53.

Andrej Bauer has recently shown that proofs and results in BISH can be translated into Weihrauch's Type-2 Effectivity framework by a realisability interpretation [5].

The principle $(\neg\neg P \Longrightarrow P)$ is equivalent to LEM even with intuitionistic logic; see Exercise 1.

For the confirmed intuitionist there is at least one other reason for rejecting MP: it contradicts Brouwer's (admittedly controversial) theory of the creating subject, an add-on to his intuitionistic mathematics. See [34] (pages 116–117) and, for a rather different view, [73].

Bishop used "subfinite" instead of "finitely enumerable". Note that a subset of a finite set need not be finitely enumerable.

In the proof of Theorem 1.4.1 we could have defined X as a set of equivalence classes under the equivalence relation $\sim$ defined by

$$(0 \sim 1) \Longleftrightarrow P,$$

but it is more in keeping with Bishop's approach to proceed by considering a special, if unusual, equality relation on X.

Some constructive mathematicians argue against even the principle of countable choice; see, for example, [76, 82].

Myhill has shown that, under the Church–Markov–Turing thesis, the power set of a singleton is uncountable: in other words, there is no recursive mapping of $\mathbb{N}$ onto that power set [74] (page 364, Theorem 3). Within BISH we may not be able to prove the uncountability of the power set of $\{0\}$, but we certainly have an unending supply of subsets S of $\{0\}$ for which we cannot decide that $S = \varnothing$ or $S = \{0\}$: given any constructively meaningful statement P, we can define a corresponding set

$$S = \{x : x = 0 \wedge P\}$$

such that $(S = \{0\} \vee S = \varnothing)$ if and only if $P \vee \neg P$.

Myhill's formal theory—based on primitive notions of "set", "function", and "natural number"—is but one of several foundational theories advocated for constructive mathematics. Others include Aczel's constructive set theory [3], Martin-Löf's type theory [69], and a largely unpublished constructive version of Morse's set theory [16] in which membership of the universe appears to correspond to being constructively defined. For more on constructive foundational theories see [6] and [88].

2

Techniques of Elementary Analysis

My deplorable mania for analysis exhausts me.

—Gustave Flaubert, Letter (August 1846)

We begin by using a form of interval arithmetic as a foundation for the construction of the real number line $\mathbb{R}$. Having discussed the elementary algebraic and order-theoretic properties of real numbers, we prove that $\mathbb{R}$ is complete in two senses: Dedekind (order) complete and Cauchy (sequentially) complete. The next step is to generalise from the reals to metric and normed spaces. A particularly important property for us is total boundedness, which plays a big part in proving the existence of suprema and infima; for that reason we need to prove that there are lots of totally bounded subsets in a totally bounded space. We also highlight the important property of locatedness for subsets of a metric space, a property that holds automatically in classical mathematics.

2.1 The Real Number Line

Constructive analysis proper, as distinct from arithmetic, begins with the real numbers, which we shall construct, choice-free, by interval arithmetic. First, though, we observe that the purely algebraic constructions of the sets $\mathbb{Z}$ of integers and $\mathbb{Q}$ of rational numbers from $\mathbb{N}$ are carried out as in classical algebra, and that the standard inequality on $\mathbb{Q}$—and hence a fortiori on $\mathbb{Z}$ regarded as a subset of $\mathbb{Q}$—is the (in this case discrete) denial inequality.

By a *real number* we mean a subset $\mathbf{x}$ of $\mathbb{Q} \times \mathbb{Q}$ such that

- $\triangleright$ for all (q, q') in $\mathbf{x}$, $q \leqslant q'$;
- $\triangleright$ for all (q, q') and (r, r') in $\mathbf{x}$, the closed intervals $[q, q']$ and $[r, r']$ in $\mathbb{Q}$ intersect in points of $\mathbb{Q}$;

▷ for each positive rational ε there exists (q, q') in $\mathbf{x}$ such that $q' - q < \varepsilon$.

The last of these properties ensures that $\mathbf{x}$ is inhabited. It also legitimises our use of expressions like "pick an element of $\mathbf{x}$": to carry out such a selection, we simply take $\varepsilon = 1$ in the third of the defining properties of the real number $\mathbf{x}$, to produce a corresponding element (q, q') of $\mathbf{x}$.

The intuition underlying our definition of "real number" is that the elements of $\mathbf{x}$ are the rational endpoints of closed intervals with one point—namely $\mathbf{x}$—in common. Any rational number q gives rise to a canonical real number

$$\mathbf{q} = \{(q, q)\}$$

with which the original rational q is identified.

Two real numbers $\mathbf{x}$ and $\mathbf{y}$ are

- *equal,* written $\mathbf{x} = \mathbf{y}$, if for all $(q, q') \in \mathbf{x}$ and all $(r, r') \in \mathbf{y}$, the intervals $[q, q']$ and $[r, r']$ in $\mathbb{Q}$ have a rational point in common;
- *unequal* (or *distinct*), written $\mathbf{x} \neq \mathbf{y}$, if there exist $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ such that the intervals $[q, q']$ and $[r, r']$ in $\mathbb{Q}$ are disjoint.

It is almost immediate that $\neq$ satisfies the defining properties of an inequality relation. Let us check that equality is an equivalence relation. It is trivial that it is reflexive and symmetric, so only transitivity has to be handled. Let $\mathbf{x} = \mathbf{y}$ and $\mathbf{y} = \mathbf{z}$, and suppose that for some $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{z}$ there are no rational points in $[q, q'] \cap [r, r']$. We may assume without loss of generality that $q' < r$. Using the third of the defining properties for a real number, choose $(s, s') \in \mathbf{y}$ such that $s' - s < r - q'$. Now, the rational interval $[s, s']$ intersects $[q, q']$ in a rational point u, and $[r, r']$ in a rational point v, so

$$r - q' \leqslant v - u \leqslant s' - s < r - q',$$

a contradiction. Hence

$$\neg \left([q, q'] \cap [r, r'] = \varnothing \right).$$

Since we are working with intervals in $\mathbb{Q}$, with the aid of two simple lemmas we can turn this around to construct a point of $[q, q'] \cap [r, r']$, and therefore complete the proof that $\mathbf{x} = \mathbf{z}$, as follows.

Lemma 2.1.1. *Let a, b, c, d be rational numbers with $a \leqslant b$ and $c \leqslant d$. Then there exists a rational number in $[a, b] \cap [c, d]$ if and only if $a \leqslant d$ and $c \leqslant b$.*

Proof. If r is a rational number in the intersection of the intervals, then $a \leqslant r \leqslant d$ and $c \leqslant r \leqslant b$, so $a \leqslant d$ and $c \leqslant b$. If, conversely, these conditions hold, then either $c < a$ and therefore $a \in [a, b] \cap [c, d]$, or else $a \leqslant c$ and $c \in [a, b] \cap [c, d]$. □

Lemma 2.1.2. *Let I, J be closed, bounded intervals in $\mathbb{Q}$ such that $\neg(I \cap J = \varnothing)$. Then there exists $r \in \mathbb{Q}$ such that $r \in I \cap J$.*

Proof. Let $I = [a, b]$ and $J = [c, d]$. If $b < c$, then $I \cap J = \varnothing$, a contradiction. Hence $c \leqslant b$. Likewise, $a \leqslant d$. It remains to apply Lemma 2.1.1. □

Taken with the equality and inequality we have defined above, the collection of real numbers forms a set: the *real line* $\mathbb{R}$.

Let $\mathbf{x}, \mathbf{y}$ be real numbers. We say that $\mathbf{x}$ is *greater than* $\mathbf{y}$, and that $\mathbf{y}$ is *less than* $\mathbf{x}$, if there exist $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ such that $r' < q$; we then write $\mathbf{x} > \mathbf{y}$ or, equivalently, $\mathbf{y} < \mathbf{x}$. On the other hand, we say that $\mathbf{x}$ is *greater than or equal to* $\mathbf{y}$, and that $\mathbf{y}$ is *less than or equal to* $\mathbf{x}$, if for all $(q, q') \in \mathbf{x}$ and all $(r, r') \in \mathbf{y}$ we have $q' \geqslant r$; we then write $\mathbf{x} \geqslant \mathbf{y}$ or, equivalently, $\mathbf{y} \leqslant \mathbf{x}$. We write, for example, $x \not> y$ to indicate that $\neg(x > y)$, and $x \nleqslant y$ to indicate that $\neg(x \leqslant y)$.

Clearly, $\mathbf{x} \geqslant \mathbf{x}$ and $\mathbf{x} \not> \mathbf{x}$. Moreover, by Lemma 2.1.1, $\mathbf{x} = \mathbf{y}$ if and only if both $\mathbf{x} \geqslant \mathbf{y}$ and $\mathbf{y} \geqslant \mathbf{x}$; and $\mathbf{x} \neq \mathbf{y}$ if and only if either $\mathbf{x} > \mathbf{y}$ or else $\mathbf{x} < \mathbf{y}$.

Lemma 2.1.3. *If $\mathbf{x} > \mathbf{y}$, then $\mathbf{y} \not> \mathbf{x}$.*

Proof. Since $\mathbf{x} > \mathbf{y}$, there exist $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ such that $r' < q$. If also $\mathbf{y} > \mathbf{x}$, then there exist $(s, s') \in \mathbf{x}$ and $(t, t') \in \mathbf{y}$ such that $s' < t$. By the defining conditions on real numbers, there exist rational numbers a, b such that $a \in [q, q'] \cap [s, s']$ and $b \in [r, r'] \cap [t, t']$. But then

$$a \leqslant s' < t \leqslant b \leqslant r' < q \leqslant a,$$

a contradiction. □

Lemma 2.1.4. $\mathbf{x} \geqslant \mathbf{y}$ *if and only if* $\mathbf{y} \not> \mathbf{x}$.

Proof. By definition, $\mathbf{x} \geqslant \mathbf{y}$ if and only if for all $(q, q') \in \mathbf{x}$ and all $(r, r') \in \mathbf{y}$ we have $q' \geqslant r$, which occurs precisely when $\neg(\mathbf{y} > \mathbf{x})$. □

Lemma 2.1.5. *If $\mathbf{x} > \mathbf{y}$, then $\mathbf{x} \geqslant \mathbf{y}$.*

Proof. By Lemma 2.1.3, $\mathbf{y} \not> \mathbf{x}$. The result follows from Lemma 2.1.4. □

Lemma 2.1.6. *If $\mathbf{x} > \mathbf{0}$, then there exists a positive integer n such that $\mathbf{x} > 1/n$.*

Proof. First pick $(q, q') \in \mathbf{x}$ such that $0 < q$. Then choose a positive integer n such that $q > 1/n$. The definition of "greater than" ensures that $\mathbf{x} > 1/n$. □

From time to time we shall revisit some of our earlier Brouwerian examples in the light of our formal definitions of real numbers and their properties.

Proposition 2.1.7. *The statement*

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R} \, (\neg \, (\mathbf{x} \geqslant \mathbf{y}) \Longrightarrow \mathbf{y} > \mathbf{x}) \tag{2.1}$$

implies Markov's principle.

Proof. Assume (2.1). Let $(a_n)_{n \geqslant 1}$ be an increasing binary sequence such that

$$\neg \forall n \, (a_n = 0) \, ,$$

and define a real number by

$$\mathbf{x} = \left\{ \left(0, \frac{1}{n} \right) : a_n = 0 \right\} \cup \left\{ \left(\frac{1}{n}, \frac{1}{n} \right) : a_n = 1 - a_{n-1} \right\} . \tag{2.2}$$

Then $\neg \, (\mathbf{0} \geqslant \mathbf{x})$: for if $0 \geqslant q$ for all $(q, q') \in \mathbf{x}$, then $a_n = 0$ for all n, a contradiction. It follows from (2.1) that $\mathbf{x} > \mathbf{0}$ and hence that there exists $(q, q') \in \mathbf{x}$ such that $q > 0$. Then $(q, q') = \left(\frac{1}{n}, \frac{1}{n} \right)$ for a (unique) n such that $a_n = 1 - a_{n-1}$. □

Proposition 2.1.8. *The relation* $>$ *is* **cotransitive***: If* $\mathbf{a} < \mathbf{b}$, *then for all* $\mathbf{x} \in \mathbb{R}$ *either* $\mathbf{a} < \mathbf{x}$ *or* $\mathbf{x} < \mathbf{b}$.

Proof. There exist $(q, q') \in \mathbf{a}$ and $(r, r') \in \mathbf{b}$ such that $q' < r$. Given a real number $\mathbf{x}$, we can find $(s, s') \in \mathbf{x}$ such that $s' - s < r - q'$. If $s' < r$, then $\mathbf{x} < \mathbf{b}$; if $s' \geqslant r$, then $q' < s$ and so $\mathbf{a} < \mathbf{x}$. □

Lemma 2.1.9. *If* $\mathbf{x} > \mathbf{y} \geqslant \mathbf{z}$ *or* $\mathbf{x} \geqslant \mathbf{y} > \mathbf{z}$, *then* $\mathbf{x} > \mathbf{z}$ *and* $\mathbf{x} \geqslant \mathbf{z}$.

Proof. Assume, for example, that $\mathbf{x} > \mathbf{y} \geqslant \mathbf{z}$. By Proposition 2.1.8, either $\mathbf{x} > \mathbf{z}$ or $\mathbf{z} > \mathbf{y}$. Since the latter is ruled out by Lemma 2.1.4, we have $\mathbf{x} > \mathbf{z}$ and therefore, by Lemma 2.1.5, $\mathbf{x} \geqslant \mathbf{z}$. □

Lemma 2.1.10. *If* $(q, q') \in \mathbf{x}$, *then* $\mathbf{q} \leqslant \mathbf{x} \leqslant \mathbf{q}'$.

Proof. By Lemma 2.1.1, for each $(r, r') \in \mathbf{x}$, since the rational intervals $[r, r']$ and $[q, q']$ intersect, $r' \geqslant q$ and $q' \geqslant r$. The desired conclusion now follows from the definition of the relation $\leqslant$. □

Lemma 2.1.11. *For each real number* $\mathbf{x}$ *there exist rational numbers* $\mathbf{q}, \mathbf{q}'$ *such that* $\mathbf{q} < \mathbf{x} < \mathbf{q}'$.

Proof. Let (r, r') be any element of $\mathbf{x}$. By Lemma 2.1.10, $r \leqslant \mathbf{x} \leqslant r'$. Choosing q, q' in $\mathbb{Q}$ with $q < r \leqslant r' < q'$, we see from the definition of the relation $<$ that $\mathbf{q} < \mathbf{x} < \mathbf{q}'$. $\square$

The following two propositions show that the classical law of trichotomy does not hold constructively, even in the weak form discussed in Proposition 2.1.13.

Proposition 2.1.12. *The statement*

$$\forall \mathbf{x} \in \mathbb{R} \ \ (\mathbf{x} \geqslant \mathbf{0} \Longrightarrow \mathbf{x} > \mathbf{0} \vee \mathbf{x} = \mathbf{0})$$

implies LPO.

Proof. Given an increasing binary sequence $(a_n)_{n \geqslant 1}$, define the real number $\mathbf{x}$ as at (2.2). It is routine to check that $\mathbf{x} \geqslant \mathbf{0}$; that if $\mathbf{x} = \mathbf{0}$, then $a_n = 0$ for all n; and that if $\mathbf{x} > \mathbf{0}$, then there exists n such that $a_n = 1$. $\square$

Proposition 2.1.13. *The statement*

$$\forall \mathbf{x} \in \mathbb{R} \ \ (\mathbf{x} \geqslant \mathbf{0} \vee \mathbf{x} \leqslant \mathbf{0})$$

implies LLPO.

Proof. Let $(a_n)_{n \geqslant 1}$ be a binary sequence with at most one term equal to 1, and define a real number by

$$\mathbf{x} = \left\{ \left(-\frac{1}{n}, \frac{1}{n} \right) : \forall k \leqslant n \, (a_k = 0) \right\} \cup \left\{ \left((-1)^n \, \frac{1}{n}, (-1)^n \, \frac{1}{n} \right) : a_n = 1 \right\}.$$

If $\mathbf{x} \geqslant \mathbf{0}$, then it is impossible that $a_n = 1$ for an odd n, so $a_n = 0$ for all odd n. Likewise, if $\mathbf{x} \leqslant \mathbf{0}$, then $a_n = 0$ for all even n. $\square$

Lemma 2.1.14. *If* $\mathbf{x} < \mathbf{y}$, *then there exists* $\mathbf{s} \in \mathbb{Q}$ *such that* $\mathbf{x} < \mathbf{s} < \mathbf{y}$.

Proof. Since $\mathbf{x} < \mathbf{y}$, there exist $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ such that $q' < r$. Then the rational number

$$\mathbf{s} = \left\{ \left(\frac{1}{2} \left(q' + r \right), \frac{1}{2} \left(q' + r \right) \right) \right\}$$

has the desired property. $\square$

The *maximum* of two real numbers $\mathbf{x}$ and $\mathbf{y}$ is the set $\max\{\mathbf{x},\mathbf{y}\}$ of all rational pairs of the form $(\max\{q,r\},\max\{q',r'\})$ where $(q,q')\in\mathbf{x}$, $(r,r')\in\mathbf{y}$, and, for example, $\max\{q,r\}$ is the maximum in $\mathbb{Q}$ of the two rational numbers q,r. The *minimum* of $\mathbf{x}$ and $\mathbf{y}$ is the set $\min\{\mathbf{x},\mathbf{y}\}$ of all rational pairs of the form $(\min\{q,r\},\min\{q',r'\})$ where $(q,q')\in\mathbf{x}$, $(r,r')\in\mathbf{y}$, and $\min\{q',r'\}$ is the minimum in $\mathbb{Q}$ of the two rational numbers q',r'. The maximum, $\max\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$, and minimum, $\min\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$, of finitely many real numbers are defined analogously. We define the *negative* of a real number $\mathbf{x}$ to be

$$^{-}\mathbf{x}=\left\{(q,q')\in\mathbb{Q}\times\mathbb{Q}:\left(^{-}q',^{-}q\right)\in\mathbf{x}\right\},$$

where, for a rational number q, the expression ^{-}q denotes the negative defined in the usual way. The *absolute value* of the real number $\mathbf{x}$ is the set

$$|\mathbf{x}|=\max\left\{^{-}\mathbf{x},\mathbf{x}\right\}.$$

It is not hard to verify such properties as the following:

- $\max\{\mathbf{x},\mathbf{y}\}$, $\min\{\mathbf{x},\mathbf{y}\}$, $^{-}\mathbf{x}$, and $|\mathbf{x}|$ are real numbers.
- $\max\{\mathbf{x},\mathbf{y}\}<\mathbf{z}$ if and only if $\mathbf{x}<\mathbf{z}$ and $\mathbf{y}<\mathbf{z}$.
- $\max\{\mathbf{x},\mathbf{y}\}>\mathbf{z}$ if and only if either $\mathbf{x}>\mathbf{z}$ or $\mathbf{y}>\mathbf{z}$.
- $\min\{\mathbf{x},\mathbf{y}\}<\mathbf{z}$ if and only if either $\mathbf{x}<\mathbf{z}$ or $\mathbf{y}<\mathbf{z}$.
- $\mathbf{x}<\mathbf{y}$ if and only if $^{-}\mathbf{y}<^{-}\mathbf{x}$.
- $\max\{\mathbf{x},\mathbf{y}\}=\ ^{-}\min\{^{-}\mathbf{x},^{-}\mathbf{y}\}$.
- $|\mathbf{x}|<\mathbf{y}$ if and only if $^{-}\mathbf{y}<\mathbf{x}<\mathbf{y}$.

In order to avoid tedious, detailed verifications of all the usual facts about real numbers, we shall often take such properties for granted.

Lemma 2.1.15. *For each real number* $\mathbf{x}$ *there exists a positive integer* n *such that* $|\mathbf{x}|<n$.

Proof. Let $(q,q')\in\mathbf{x}$, and choose a positive integer n such that

$$\max\{|q|,|q'|\}<n.$$

Using rational arithmetic, we see that $-n<q\leqslant q'<n$; whence, by the definition of the relation $<$, we have $-n<\mathbf{x}<n$ and therefore $|\mathbf{x}|<n$. □

There are two common misconceptions about the constructive properties of the real line $\mathbb{R}$: that it is countable and that it is not complete. The following lemma enables us to dispose of the first of these misconceptions.

Lemma 2.1.16. *Let $I = [0,1]$, and for $i = 1,2,3$ let $J_i = \left[\frac{i-1}{3}, \frac{i}{3}\right]$. Then for each $\mathbf{x} \in \mathbb{R}$ there exists $i \in \{1,3\}$ such that $\mathbf{x} \neq \mathbf{y}$ for all $\mathbf{y} \in J_i$.*

Proof. This is immediate, since the relation $<$ is transitive and, by Proposition 2.1.8, either $\mathbf{x} > 1/3$ or $\mathbf{x} < 2/3$. □

We refer to the interval J_1 in Lemma 2.1.16 as the *leftmost third* of I, and to J_3 as the *rightmost third.* The uncountability of $\mathbb{R}$ is a simple consequence of the following version of *Cantor's theorem.*

Theorem 2.1.17. *If $(\mathbf{a}_n)_{n\geqslant 1}$ is a sequence of real numbers, then there exists $\mathbf{x} \in [0,1]$ such that $\mathbf{x} \neq \mathbf{a}_n$ for each n.*

Proof. We construct inductively a sequence $((q_n, q_n'))_{n\geqslant 1}$ in $\mathbb{Q} \times \mathbb{Q}$ such that for all $n \geqslant 1$,

- $0 < q_n' - q_n = 3^{-n}$,
- $[q_n, q_n'] \subset \left[q_{n-1}, q_{n-1}'\right]$, and
- $\mathbf{a}_n \neq \mathbf{y}$ for each $\mathbf{y} \in [q_n, q_n']$.

For convenience we take $q_0 = 0$ and $q_0' = 1$. Suppose that we have already constructed the elements (q_k, q_k') $(0 \leqslant k \leqslant n-1)$ of $\mathbb{Q} \times \mathbb{Q}$ such that the applicable properties hold. By Lemma 2.1.16, either $\mathbf{a}_n \neq \mathbf{y}$ for all $\mathbf{y}$ in the leftmost third of the interval $\left[q_{n-1}, q_{n-1}'\right]$, or else $\mathbf{a}_n \neq \mathbf{y}$ for all $\mathbf{y}$ in the rightmost third. In either case we take q_n and q_n' to be respectively the left and right endpoints of the third of $\left[q_{n-1}, q_{n-1}'\right]$ in question. It is clear that the pair (q_n, q_n') satisfies the desired properties, and hence that our inductive construction is complete. It is equally clear that

$$\mathbf{x} = \{(q_n, q_n') : n \in \mathbb{N}\}$$

is a real number. Finally, for each positive integer n we have $\mathbf{x} \in [q_n, q_n']$, by Lemma 2.1.10, and therefore $\mathbf{x} \neq \mathbf{a}_n$. □

In this proof we used the principle of dependent choice for the first time and in the informal manner typical of the analyst. So it may be a good idea just once to show how this informal way of invoking dependent choice can be made more rigorous and explicit. Accordingly, let

$$X = \{(n, q, q') : n \in \mathbb{N}, q \in \mathbb{Q}, q' \in \mathbb{Q}, q < q'\}.$$

Let P be the set of all pairs $((n,q,q'), (m,r,r')) \in X \times X$ such that $m = n+1$, $[r, r'] \subset [q, q']$, $r' - r = \frac{1}{3}(q' - q)$, and $\mathbf{a}_n \neq \mathbf{y}$ for each $\mathbf{y} \in [r, r']$. Applying the principle of dependent choice, construct a mapping $f : \mathbb{N} \longrightarrow X$ such that $f(0) = (1, 0, 1)$ and $(f(n), f(n+1)) \in P$ for each n. Writing

$$f(n-1) = \left(n, q_{n-1}, q'_{n-1}\right)$$

for each positive integer n, we see that

$$[q_n, q'_n] \subset \left[q_{n-1}, q'_{n-1}\right], \qquad q'_n - q_n = \frac{1}{3}\left(q'_{n-1} - q_{n-1}\right),$$

and $\mathbf{a}_n \neq \mathbf{y}$ for each $\mathbf{y} \in [q_n, q'_n]$, as we required.

We now seek a constructive analogue of the classical least-upper-bound principle. Let S be a set of real numbers. We say that a real number $\mathbf{b}$ is an *upper bound* of/for S if $\mathbf{s} \leqslant \mathbf{b}$ for all $\mathbf{s} \in S$; and that $\mathbf{b}$ is the (perforce unique) *least upper bound* of S if it is an upper bound and for each $\mathbf{x} < \mathbf{b}$ there exists $\mathbf{s} \in S$ such that $\mathbf{x} < \mathbf{s}$. In the latter event we also call $\mathbf{b}$ the *supremum* of S and we denote it by $\sup S$.

The classical least-upper-bound principle implies the law of excluded middle. To see this, let P be a syntactically well-formed proposition, and consider

$$S = \{0\} \cup \{x \in \mathbb{R} : x = 1 \wedge P\},$$

which contains 0 and is bounded above by 1. If $\sup S$ exists, then it is either greater than 0 or less than 1. In the first case, there exists $x \in S$ such that $x > 0$; whence $x = 1$ and P holds. In the other case we must have $\neg P$.

We define a set S of real numbers to be *upper order located* if for all rational numbers a, b with $a < b$, either $\mathbf{x} \leqslant b$ for all $\mathbf{x}$ in S or else there exists $\mathbf{x} \in S$ with $a < \mathbf{x}$. We can now state the *constructive least-upper-bound principle,* which is also known as the *Dedekind (order) completeness of* $\mathbb{R}$.

Theorem 2.1.18. *Let S be an inhabited set of real numbers that is both bounded above and upper order located. Then the least upper bound of S exists.*

Proof. Let B be the set of upper bounds for S, and define

$$\boldsymbol{\xi} = \{(q, q') \in \mathbb{Q} \times \mathbb{Q} : \exists \mathbf{s} \in S \; \exists \mathbf{b} \in B \, (q \leqslant \mathbf{s} \leqslant \mathbf{b} \leqslant q')\}.$$

Taken with Lemmas 2.1.10 and 2.1.9, the hypotheses ensure that ξ is inhabited. If (q, q') and (r, r') belong to $\boldsymbol{\xi}$, then there exist $\mathbf{s}_1, \mathbf{s}_2 \in S$ and $\mathbf{b}_1, \mathbf{b}_2 \in B$ such that $q \leqslant \mathbf{s}_1 \leqslant \mathbf{b}_1 \leqslant q'$ and $r \leqslant \mathbf{s}_2 \leqslant \mathbf{b}_2 \leqslant r'$. Then $\mathbf{s}_1 \leqslant \mathbf{b}_2$, so $q \leqslant \mathbf{b}_2 \leqslant r'$; similarly, $r \leqslant q'$. By Lemma 2.1.1, there exists a rational number in $[q, q'] \cap [r, r']$. To complete the proof that $\boldsymbol{\xi}$ is a real number, we show that for each rational $\varepsilon > 0$ there exists $(q, q') \in \boldsymbol{\xi}$ with $q' - q < \varepsilon$. To this end, fix (a, a') in $\boldsymbol{\xi}$. If $a = a'$, then there is nothing to prove; so we may assume that $a < a'$. Construct rational numbers $a_0 = a < a_1 < a_2 < \cdots < a_n = a'$ such that $a_i - a_{i-1} < \varepsilon/2$ for $1 \leqslant i \leqslant n$. Since S is upper order located, either $a_2 \in B$ or else $a_1 < \mathbf{s}$ for some element $\mathbf{s}$ of S. In the first case, $(a_0, a_2) \in \boldsymbol{\xi}$ and $a_2 - a_0 < \varepsilon$. In the second, either $a_3 \in B$

and therefore $(a_1, a_3) \in \boldsymbol{\xi}$ and $a_3 - a_1 < \varepsilon$; or else $a_2 < \mathbf{s}$ for some element $\mathbf{s}$ of X. Carrying on in this way, since $a_n \in B$ we can be sure of finding $k < n-1$ such that $(a_k, a_{k+2}) \in \boldsymbol{\xi}$ and $a_{k+2} - a_k < \varepsilon$. Thus $\boldsymbol{\xi}$ is indeed a real number.

To show that ξ is an upper bound for S, consider any (q, q') in ξ and any $\mathbf{s}$ in S. There exists $\mathbf{b} \in B$ such that $q \leqslant \mathbf{b} \leqslant q'$. For any (r, r') in $\mathbf{s}$ we have $r \leqslant \mathbf{s}$, by Lemma 2.1.10, and therefore $r \leqslant \mathbf{b}$; whence $r \leqslant q'$. It follows that $\mathbf{s} \leqslant \boldsymbol{\xi}$.

Finally, if $\mathbf{x} < \boldsymbol{\xi}$, then we can find $(q, q') \in \boldsymbol{\xi}$ and $(r, r') \in \mathbf{x}$ such that $q > r'$. It follows from Lemma 2.1.10 that $\mathbf{x} \leqslant r' < q$. By definition of $\boldsymbol{\xi}$, there exists $\mathbf{s} \in S$ with $q \leqslant \mathbf{s}$; whence $\mathbf{x} < \mathbf{s}$, by Lemma 2.1.9. This completes the proof that $\boldsymbol{\xi}$ is the least upper bound for S. □

Let S be a set of real numbers. We say that a real number $\mathbf{b}$ is a *lower bound* of/for S if $\mathbf{b} \leqslant \mathbf{s}$ for all $\mathbf{s} \in S$; and that $\mathbf{b}$ is the (perforce unique) *greatest lower bound* of S if it is a lower bound and for each $\mathbf{x} > \mathbf{b}$ there exists $\mathbf{s} \in S$ such that $\mathbf{x} > \mathbf{s}$. In the latter event we also call $\mathbf{b}$ the *infimum* of S and we denote it by $\inf S$.

Corollary 2.1.19. *Let S be an inhabited set of real numbers that is bounded below and is* **lower order located** *in the following sense: for all rational numbers a, b with $a < b$, either a is a lower bound for S or else there exists $\mathbf{s} \in S$ with $\mathbf{s} < b$. Then the greatest lower bound of S exists.*

Proof. Apply Theorem 2.1.18 to the set

$$T = \left\{ {}^{-}\mathbf{s} : \mathbf{s} \in S \right\},$$

which is inhabited, bounded above, and upper order located, to construct its supremum $\mathbf{b}$. Then ${}^{-}\mathbf{b}$ is the infimum of S. □

We now want to introduce the arithmetic operations on real numbers. Given real numbers $\mathbf{x}$ and $\mathbf{y}$, we need to find the appropriate ways of combining the rational pairs (representing intervals) that constitute the real number $\mathbf{x}$ with those that constitute $\mathbf{y}$, in order to create the rational pairs that represent $\mathbf{x} \circ \mathbf{y}$, where $\circ$ stands for any of the operations $+, -, \times, \div$. We begin with the easy definitions of $+$ and $-$, for the moment leaving aside the more complicated ones for $\times$ and $\div$.

We define the *sum* $\mathbf{x} + \mathbf{y}$ and *difference* $\mathbf{x} - \mathbf{y}$ of the real numbers $\mathbf{x}, \mathbf{y}$ to be, respectively,

$$\mathbf{x} + \mathbf{y} = \{(s, s') : \exists\, (q, q') \in \mathbf{x}\; \exists\, (r, r') \in \mathbf{y}\, (s = q + r \wedge s' = q' + r')\},$$
$$\mathbf{x} - \mathbf{y} = \{(s, s') : \exists\, (q, q') \in \mathbf{x}\; \exists\, (r, r') \in \mathbf{y}\, (s = q - r \wedge s' = q' - r')\}.$$

Let us verify, for example, that $\mathbf{x}+\mathbf{y}$ is a real number. Let $(q_1, q_1') \in \mathbf{x}$ and $(r_1, r_1') \in \mathbf{y}$. Certainly, $q_1 + r_1 \leqslant q_1' + r_1'$. Moreover, given a positive rational number ε, we can arrange that $q_1' - q_1 < \varepsilon/2$ and $r_1' - r_1 < \varepsilon/2$, so $(q_1 + r_1') - (q_1 + r_1) < \varepsilon$. It remains to show that if also $(q_2, q_2') \in \mathbf{x}$ and $(r_2, r_2') \in \mathbf{y}$, then the intervals $[q_i + r_i, q_i' + r_i']$ $(i = 1, 2)$ in $\mathbb{Q}$ intersect. This is easy: there exist rational numbers ξ, η such that $q_i \leqslant \xi \leqslant q_i'$ and $r_i \leqslant \eta \leqslant r_i'$ $(i = 1, 2)$; whence

$$q_i + r_i \leqslant \xi + \eta \leqslant q_i' + r_i' \quad (i = 1, 2).$$

Thus $\mathbf{x} + \mathbf{y}$ is a real number.

Note that if $\mathbf{x} < \mathbf{y}$, then $\mathbf{x} + \mathbf{z} < \mathbf{y} + \mathbf{z}$ for all real $\mathbf{z}$; and that $\mathbf{x} - \mathbf{y} = \mathbf{x} + ({}^{-}\mathbf{y})$. In view of the latter, we adopt the normal convention of writing $-\mathbf{x}$ instead of ${}^{-}\mathbf{x}$.

Lemma 2.1.20. *For each real number* $\mathbf{x}$ *there exists a positive integer* N *such that* $\max\{|q|, |q'|\} < N$ *whenever* $(q, q') \in \mathbf{x}$ *and* $q' - q < 1$.

Proof. With n as in Lemma 2.1.15, set $N = n + 1$. If $(q, q') \in \mathbf{x}$ and $q' - q < 1$, then, using Lemma 2.1.10, we obtain

$$q \leqslant \mathbf{x} \leqslant |\mathbf{x}| < n < N$$

and therefore

$$q' < q + 1 < n + 1 = N.$$

Since $(-q', -q) \in -\mathbf{x}$ and $|-\mathbf{x}| = |\mathbf{x}|$, we likewise have $-q' < N$ and $-q < N$. Hence $|q| < N$ and $|q'| < N$, so $\max\{|q|, |q'|\} < N$. □

We next define the *product* $\mathbf{x} \times \mathbf{y}$, also written $\mathbf{x} \cdot \mathbf{y}$ or $\mathbf{xy}$, to be the set of all rational pairs (s, s') such that there exist $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ with

$$s = \min\{qr, qr', q'r, q'r'\}, \; s' = \max\{qr, qr', q'r, q'r'\}.$$

Certainly, $s \leqslant s'$. Using Lemma 2.1.20, compute a positive integer N such that if $(q, q') \in \mathbf{x}$ and $q' - q < 1$, then $\max\{|q|, |q'|\} < N$, and such that if $(r, r') \in \mathbf{y}$ and $r' - r < 1$, then $\max\{|r|, |r'|\} < N$. Given a rational ε with $0 < \varepsilon < 1$, choose $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ such that $q' - q < \varepsilon/2N$ and $r' - r < \varepsilon/2N$. Routine rational arithmetic calculations show that

$$s' - s < (q' - q)\max\{|r|, |r'|\} + (r' - r)\max\{|q|, |q'|\} < \varepsilon.$$

For $i = 1, 2$ let $(q_i, q_i') \in \mathbf{x}$, $(r_i, r_i') \in \mathbf{y}$, and

$$s_i = \min\{q_i r_i, q_i r_i', q_i' r_i, q_i' r_i'\}, \; s_i' = \max\{q_i r_i, q_i r_i', q_i' r_i, q_i' r_i'\}.$$

Pick ξ in $[q_1, q_1'] \cap [q_2, q_2']$ and η in $[r_1, r_1'] \cap [r_2, r_2']$; it is easy to verify that $\xi\eta \in [s_1, s_1'] \cap [s_2, s_2']$. This completes the proof that $\mathbf{xy}$ is a real number.

When dealing with division, we consider two real numbers $\mathbf{x}, \mathbf{y}$ with $\mathbf{y} \neq \mathbf{0}$. In the case $\mathbf{y} > \mathbf{0}$ we construct a rational pair (s, s') in the *quotient* $\mathbf{x}/\mathbf{y}$ (also written

$\frac{\mathbf{x}}{\mathbf{y}}$) as follows. We pick $(q,q') \in \mathbf{x}$ and $(r,r') \in \mathbf{y}$ such that $r > 0$. If $q \geqslant 0$, we set $s = q/r', s' = q'/r$; if $q < 0$, we set $s = q/r, s' = q'/r'$. In the case $\mathbf{y} < \mathbf{0}$ we define $\mathbf{x}/\mathbf{y}$ to be $-((-\mathbf{x})/\mathbf{y})$. We leave as an exercise the details that $\mathbf{x}/\mathbf{y}$ is indeed a real number, and that $\mathbf{x}(1/\mathbf{x}) = \mathbf{1}$.

As classically, we say that a sequence $(\mathbf{x}_n)_{n\geqslant 1}$ of real numbers *converges* to a real number $\mathbf{x}_\infty$, called the *limit* of the sequence, if

$$\forall \varepsilon > 0 \; \exists N \; \forall n \geqslant N \left(|\mathbf{x}_\infty - \mathbf{x}_n| < \varepsilon \right).$$

We then write

$$\mathbf{x}_n \longrightarrow \mathbf{x}_\infty \text{ as } n \longrightarrow \infty$$

or

$$\mathbf{x}_\infty = \lim_{n\to\infty} \mathbf{x}_n.$$

The uniqueness of the limit, and the basic algebraic properties of sequences of real numbers, will be assumed without proof, since the proofs are virtually the same as their classical counterparts. However, we should note that the *monotone convergence theorem for sequences,*

> *Every increasing sequence in $\mathbb{R}$ that is bounded above (that is, whose terms form a set that is bounded above) converges,*

implies LPO.

By a *Cauchy sequence* of real numbers we mean a sequence $(\mathbf{x}_n)_{n\geqslant 1}$ such that

$$\forall \varepsilon > 0 \, \exists N \, \forall m,n \geqslant N \; \left(|\mathbf{x}_m - \mathbf{x}_n| < \varepsilon \right).$$

Every convergent sequence of real numbers is a Cauchy sequence. We now counter the second major misconception about the constructive nature of $\mathbb{R}$ by establishing its so-called *(Cauchy) completeness.*

Theorem 2.1.21. *Every Cauchy sequence of real numbers converges to a real number.*

Proof. Let $(\mathbf{x}_n)_{n\geqslant 1}$ be a Cauchy sequence of real numbers, and, using countable choice, compute a function $k \rightsquigarrow n_k$ from $\mathbb{N}^+$ to $\mathbb{N}^+$ such that

$$\forall k \; \forall m,n \geqslant n_k \left(|\mathbf{x}_m - \mathbf{x}_n| < 2^{-k} \right).$$

Again using countable choice, construct a sequence $((q_k, q_k'))_{k\geqslant 1}$ such that

$$\forall k \left((q_k, q_k') \in \mathbf{x}_{n_k} \wedge q_k' - q_k < 2^{-k} \right).$$

Define rational numbers

$$r_k = q_k - 2^{-k}, \quad r_k' = q_k' + 2^{-k}.$$

Then for all $n \geqslant n_k$,

$$r_k \leqslant \mathbf{x}_{n_k} - 2^{-k} < \mathbf{x}_n < \mathbf{x}_{n_k} + 2^{-k} \leqslant r'_k. \tag{2.3}$$

It follows that for all $j \geqslant k$,

$$\mathbf{x}_{n_j} \in \left[r_j, r'_j\right] \cap \left[r_k, r'_k\right].$$

Since $r'_k - r_k < 2^{-k+2} \longrightarrow 0$ as $k \longrightarrow \infty$, we conclude that

$$\mathbf{x}_\infty = \left\{(r_k, r'_k) : k \geqslant 1\right\}$$

is a real number. From (2.3) we have $\mathbf{x}_n \in [r_k, r'_k]$ for all $n \geqslant n_k$. It follows that

$$\forall k \, \forall n \geqslant n_k \left(|\mathbf{x}_n - \mathbf{x}_\infty| \leqslant r'_k - r_k < 2^{-k+2}\right).$$

Noting Lemma 2.1.6, we see that $\mathbf{x}_n \longrightarrow \mathbf{x}_\infty$ as $n \longrightarrow \infty$. □

We shall see in later chapters that completeness is often used in constructive mathematics to prove propositions that are immediate consequences of omniscience principles.

For the remainder of this book we shall drop the use of boldface type to denote real numbers; it has served its purpose to signal a distinction between a rational number and a real number (a set of special pairs of rational numbers). We shall also assume, without further comment, basic properties of real numbers—for example, $x^2 \geqslant 0$ for all real x—that can easily be deduced from the foregoing results.

The *complex plane* $\mathbb{C}$ consists of all *complex numbers*—ordered pairs (x, y) of real numbers—with addition and multiplication defined by

$$\begin{aligned}(x, y) + (x', y') &= (x + x', y + y'), \\ (x, y) \times (x', y') &= (xx' - yy', xy' + x'y).\end{aligned}$$

The equality and inequality on $\mathbb{C}$ are defined by

$$\begin{aligned}(x, y) = (x', y') &\Longleftrightarrow x = x' \wedge y = y' \\ (x, y) \neq (x', y') &\Longleftrightarrow x \neq x' \vee y \neq y'.\end{aligned}$$

We embed $\mathbb{R}$ as a subset of $\mathbb{C}$ in the usual way by identifying the real number x with the complex number $(x, 0)$. The pair $\mathrm{i} = (0, 1)$ then has the special property that $\mathrm{i}^2 = -1$; and every complex number $z = (x, y)$ can be written in the form $x + \mathrm{i}y$, with *real part* $\operatorname{Re} z = x$ and *imaginary part* $\operatorname{Im} z = y$. The *complex conjugate* of $z = x + \mathrm{i}y$ is the number $x - \mathrm{i}y$, denoted by z^*. We shall assume basic definitions and properties of $\mathbb{C}$, such as its completeness, as they are needed. The same goes for the Euclidean spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, which are now defined in the standard ways.

2.2 Metric Spaces

Since we are assuming some familiarity with the classical theory of metric spaces, in the following we shall emphasise differences between the classical and the constructive theory, as well as those constructive properties, such as locatedness, that play no role in classical analysis. We normally denote the metric on a set X by ρ.

The definitions of such notions as *metric*, *(metric) subspace, ball, interior,* and *cluster point* are exactly as in classical analysis; see, for example, [47]. The interior of a subset S of a metric space X is denoted by S° and is a subset of S; if $S = S^\circ$, then S is said to be *open.* The set $\overline{S}$ of cluster points of S is called the *closure* of S (in X), and contains S; if $S = \overline{S}$, then S is said to be *closed* in X. If $\overline{S} = X$, then S is *dense* in in X. This happens if and only if for each $x \in X$ and each $\varepsilon > 0$ there exists $s \in S$ such that $\rho\,(x, s) < \varepsilon$.

We denote the open and closed balls in X with centre a and radius $r > 0$ by $B(a, r)$ and $\overline{B}(a, r)$ respectively. The *inequality* on a metric space is defined by

$$x \neq y \Longleftrightarrow \rho(x, y) > 0.$$

If $X_1, \dots, X_n$ are metric spaces, with metrics $\rho_1, \dots, \rho_n$ respectively, then the Cartesian product $X = X_1 \times \cdots \times X_n$ is a metric space relative to the *product metric*

$$\rho = \rho_1 + \cdots + \rho_n.$$

We then call (X, ρ) the *product of the metric spaces* $X_1, \dots, X_n$.

As in the classical theory, both X and $\varnothing$ are open, arbitrary unions of open sets are open, and finite intersections of open sets are open; in other words, the open sets form a topology on X. Likewise, both X and $\varnothing$ are closed, and arbitrary intersections of closed sets are closed. But we cannot prove that the union of two closed sets is closed, even in the case $X = \mathbb{R}$ and the two closed sets are closed intervals: both the intervals $[-1, 0]$ and $[0, 1]$ are closed in $\mathbb{R}$, and their union is dense in $[-1, 1]$; but if that union is closed, then we have

$$\forall x \in \mathbb{R}\ \ (x \leqslant 0 \vee x \geqslant 0)\,,$$

a proposition equivalent to LLPO.

Proposition 2.2.1. *The logical complement of an open subset S of X is closed in X, and coincides with the complement of S.*

Proof. Given a cluster point a of $\neg S$ and any point x of S, choose, in turn, $r > 0$ such that $B(x, r) \subset S$, and $y \in \neg S$ such that $\rho(a, y) < r/2$. If $\rho(a, x) < r/2$, then

$$\rho(x, y) \leqslant \rho(a, x) + \rho(a, y) < r,$$

so $y \in S$, which is absurd. Hence, by Proposition 2.1.8, $\rho(a, x) > 0$. Since $x \in S$ is arbitrary, it follows that $a \in {\sim}S \subset \neg S$. Thus $\neg S$ is closed, and $\neg S \subset {\sim}S \subset \neg S$;

whence $\neg S = {\sim}S$. $\square$

In contrast to the classical situation, we cannot prove constructively that the complement of a closed subset of $\mathbb{R}$ is open (Exercise 9).

The classical definitions of *convergent sequence, Cauchy sequence*, and *complete metric space* carry over unchanged into the constructive setting, as they did in the special case of the metric space $\mathbb{R}$. A complete subset of a metric space X is closed in X; and a closed subset of a complete space is complete. Moreover, the product of finitely many complete spaces is complete.

We now introduce the first of two major themes of metric space theory. Given $\varepsilon > 0$, by an *ε-approximation* to a subset S of a metric space X we mean an inhabited subset T of S such that for each $s \in S$ there exists $t \in T$ with $\rho(s,t) < \varepsilon$. If for each $\varepsilon > 0$ there exists a finite ε-approximation to S, then we say that S is *totally bounded.*

The closure of a totally bounded subset of X is totally bounded. If a subset S of X contains a dense totally bounded set, then S itself is totally bounded. The product of finitely many totally bounded spaces is totally bounded.

There are at least two reasons why total boundedness is so important in constructive analysis. The first is that with its help we can compute suprema and infima in many important situations. The second is that, coupled with completeness, total boundedness gives the only one of three classically equivalent notions of compactness that can be used in constructive analysis.

The next lemma enables us to replace "finite" by the weaker "finitely enumerable" in the definition of total boundedness.

Lemma 2.2.2. *Let ε be a positive number, and X a metric space with a finitely enumerable ε-approximation. Then X has a finite η-approximation for each $\eta > \varepsilon$.*

Proof. Let $\{x_1, \ldots, x_n\}$ be an ε-approximation to X, and let $\eta > \varepsilon$. We may assume that $n > 1$. Either $\rho(x_i, x_j) > 0$ whenever $i \neq j$, or $\rho(x_i, x_j) < \eta - \varepsilon$ for some pair of distinct indices i, j. In the former case, $\{x_1, \ldots, x_n\}$ is a finite η-approximation. In the latter case we can delete x_j from $\{x_1, \ldots, x_n\}$, to obtain a finitely enumerable ε'-approximation with $\varepsilon' = \varepsilon + \rho(x_i, x_j) < \eta$. Applying this argument at most $n-1$ times, we arrive at a finite η-approximation to X. $\square$

Corollary 2.2.3. *A metric space X is totally bounded if and only if for each $\varepsilon > 0$ there exists a finitely enumerable ε-approximation to X.*

Proof. This is an immediate consequence of Lemma 2.2.2. $\square$

A metric space X is *separable* if it has a countable dense subset S; an enumeration $(x_n)_{n \geqslant 1}$ of S is then called a *dense sequence* in X. The real line is separable, since, by Lemma 2.1.10 and the definition of "real number", it has $\mathbb{Q}$ as a countable dense subset.

Proposition 2.2.4. *A totally bounded metric space is separable.*

Proof. If, for each positive integer n, F_n is a finite $1/n$-approximation to the metric space X, then $\bigcup_{n \geqslant 1} F_n$ is a countable dense subset of X. □

We now have what is among the most useful of all results about totally bounded sets.

Proposition 2.2.5. *If S is a totally bounded subset of $\mathbb{R}$, then $\sup S$ and $\inf S$ exist.*

Proof. We first consider the case where $S = \{x_1, \ldots, x_n\}$ is finitely enumerable. Given real numbers a, b with $a < b$, we apply Proposition 2.1.8 repeatedly, to prove that either $x_k < b$ for each k or else there exists j such that $x_j > a$. It follows from the constructive least-upper-bound principle that $\sup S$ exists.

Now consider the general case. Again let a, b be real numbers with $a < b$, but this time write $\varepsilon = (b - a)/2$ and construct a finite ε-approximation $\{x_1, \ldots, x_n\}$ to S. By the first part of the proof,

$$\sigma = \sup \{x_1, \ldots, x_n\}$$

exists. By Proposition 2.1.8, either $\sigma > a$ or $\sigma < a + \varepsilon$. In the first case there exists j such that $x_j > a$. In the second, consider any $x \in S$. Choosing j such that $|x - x_j| < \varepsilon$, we have

$$x \leqslant x_j + |x - x_j| < \sigma + \varepsilon < a + 2\varepsilon = b.$$

So in this case, b is an upper bound for S. It follows from Theorem 2.1.18 that $\sup S$ exists. Similar arguments show that $\inf S$ exists. □

A mapping $f : X \longrightarrow Y$ between metric spaces is

- *sequentially continuous at $x \in X$* if for each sequence $(x_n)_{n \geqslant 1}$ that converges in X to x, the sequence $(f(x_n))_{n \geqslant 1}$ converges in Y to $f(x)$;
- *sequentially continuous* if it is sequentially continuous at each point of X;
- *continuous at $x \in X$* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\rho(f(x), f(x')) < \varepsilon$ for all $x' \in X$ with $\rho(x, x') < \delta$;
- *continuous* if it is continuous at each point of X.

Although sequential continuity is classically equivalent to continuity, the most we can say constructively is that the latter notion implies the former.

For constructive practice it is frequently necessary to assume the following stronger continuity property. We say that $f : X \longrightarrow Y$ is *uniformly continuous* (on X) if for each $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, x' \in X$, if $\rho(x, x') < \delta$, then $\rho(f(x), f(x')) < \varepsilon$. Clearly, uniform continuity implies continuity. Although in CLASS and in INT every continuous mapping from the bounded closed interval $[0, 1]$ into $\mathbb{R}$ is uniformly continuous, since this is not the case in RUSS we cannot expect to prove it in BISH; see Chapter 3 of [34].

Proposition 2.2.6. *If X is a totally bounded space, and f a uniformly continuous mapping of X into a metric space, then $f(X)$ is totally bounded.*

Proof. Given $\varepsilon > 0$, compute $\delta > 0$ such that if $x, y \in X$ and $\rho(x, y) < \delta$, then $\rho(f(x), f(y)) < \varepsilon$. There exists a finite δ-approximation $\{x_1, \ldots, x_n\}$ to X. For each $x \in X$ pick i such that $\rho(x, x_i) < \delta$. Then $\rho(f(x), f(x_i)) < \varepsilon$. Hence $\{f(x_i), \ldots, f(x_n)\}$ is a finitely enumerable ε-approximation to $f(X)$. □

Corollary 2.2.7. *Let f be a uniformly continuous mapping of a totally bounded space into $\mathbb{R}$. Then* $\operatorname{ran}(f)$ *is bounded, and the* **supremum**

$$\sup_{x \in X} f(x) = \sup \{y : y \in \operatorname{ran}(f)\}$$

and **infimum**

$$\inf_{x \in X} f(x) = \inf \{y : y \in \operatorname{ran}(f)\}$$

of f exist.

Proof. This is an immediate consequence of Propositions 2.2.6 and 2.2.5. □

Corollary 2.2.8. *If X is a totally bounded metric space, then its* **diameter***,*

$$\operatorname{diam}(X) = \sup \{\rho(x, y) : x, y \in X\},$$

exists.

Proof. Since the mapping ρ on the totally bounded product space $X \times X$ is uniformly continuous, the desired conclusion follows from Corollary 2.2.7. □

We now introduce a second major theme in constructive metric space theory. An inhabited subset S of a metric space X is *located* in X if for each $x \in X$ the *distance*

$$\rho(x, S) = \inf\{\rho(x, s) : s \in S\}$$

exists.

We cannot prove the proposition "every inhabited subset of $\mathbb{R}$ is located" constructively, since it implies LEM. To see this, let P be any syntactically correct proposition, and

$$S = \{0\} \cup \{x \in \mathbb{R} : (x = 1) \wedge P\}.$$

If S is located, then either $\rho(1, S) > 0$ or $\rho(1, S) < 1$. In the first case we have $\neg P$. In the second, choosing $s \in S$ such that $\rho(1, s) < 1$, we see that $s \notin \{0\}$, so $s = 1$ and P holds.

It is hard to overemphasise the importance of locatedness in constructive analysis. To illustrate, we anticipate later sections and chapters by pointing out that the norm of a bounded linear functional on a normed space exists if and only if the kernel of the functional is located, and that locatedness plays a vital role in the separation and Hahn–Banach extension theorems. It is only when we work constructively that the significance of locatedness (a property that holds automatically under classical logic) for the proofs of many existence theorems is revealed.

A subset S of a metric space is located if and only if its closure is located; in fact, for each $x \in X$ we have

$$\rho(x, S) = \rho(x, \overline{S})$$

if either side of this equation exists. Note that even if a subset S of a metric space X is located, we may not be able to prove that its *metric complement*

$$-S = \{x \in X : \rho(x, S) > 0\}$$

is located (Exercise 11).

Proposition 2.2.9. *A totally bounded subset of a metric space is located.*

Proof. Let S be totally bounded in the metric space X. For each $x \in X$ the mapping $s \rightsquigarrow \rho(x, s)$ is uniformly continuous on S. It follows from Corollary 2.2.7 that $\inf\{\rho(x, s) : s \in S\}$ exists. □

The next two proofs use a technique that will reappear from time to time. We have a finitely enumerable set $\{x_1, \ldots, x_n\}$ of a metric space X, a located subset S of X, and positive numbers a, b with $a < b$. We write $\{1, \ldots, n\}$ as a union of subsets P, Q such that

$$i \in P \Longrightarrow \rho(x_i, S) < b,$$
$$i \in Q \Longrightarrow \rho(x_i, S) > a.$$

By concentrating on one of those sets, we are able to reach our desired conclusion.

Proposition 2.2.10. *A located subset of a totally bounded metric space is totally bounded.*

Proof. Let S be a located subset of a totally bounded metric space X, and let $\varepsilon > 0$. There exists a finite $\varepsilon/3$-approximation $\{x_1, \ldots, x_n\}$ to X. Using Proposition 2.1.8, write $\{1, \ldots, n\}$ as a union of two sets P and Q, where $\rho(x_i, S) < 2\varepsilon/3$ if $i \in P$, and $\rho(x_i, S) > \varepsilon/3$ if $i \in Q$. For each $i \in P$ there exists $s_i \in S$ such that $\rho(x_i, s_i) < 2\varepsilon/3$. Given $s \in S$, choose i such that $\rho(s, x_i) < \varepsilon/3$. Then $\rho(x_i, S) < \varepsilon/3$, so $i \notin Q$; whence $i \in P$ and therefore

$$\rho(s, s_i) \leqslant \rho(s, x_i) + \rho(x_i, s_i) < \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} = \varepsilon.$$

Thus $\{s_i : i \in P\}$ is a finitely enumerable ε-approximation to S. □

The next theorem will provide us with a rich supply of totally bounded, and hence located, subsets of a totally bounded space.

Proposition 2.2.11. *Let X be a totally bounded space, x_0 a point of X, and r a positive number. Then there exists a closed, totally bounded subset K of X such that $B(x_0, r) \subset K \subset \overline{B}(x_0, 8r)$.*

Proof. With $F_1 = \{x_0\}$, we construct inductively a sequence $(F_n)_{n \geqslant 1}$ of finitely enumerable subsets of X such that

(a) $\rho(x, F_n) < 2^{-n+1} r$ for each x in $B(x_0, r)$,

(b) $\rho(x, F_n) < 2^{-n+3} r$ for each x in F_{n+1}.

To this end, assume that $F_1, \ldots, F_n$ have been constructed with the appropriate properties, and let $\{x_1, \ldots, x_N\}$ be a $2^{-n}r$-approximation to X. Write $\{1, \ldots, N\}$ as a union of subsets A and B such that

$$\begin{aligned} \rho(x_i, F_n) &< 2^{-n+3} r \quad \text{if } i \in A, \\ \rho(x_i, F_n) &> 2^{-n+2} r \quad \text{if } i \in B. \end{aligned}$$

Then

$$F_{n+1} = \{x_i : i \in A\}$$

clearly satisfies the appropriate instance of (b). Let x be any point of $B(x_0, r)$. By the induction hypothesis, there exists y in F_n with $\rho(x, y) < 2^{-n+1} r$. Choosing i in $\{1, \ldots, N\}$ such that $\rho(x, x_i) < 2^{-n} r$, we have

$$\rho(x_i, F_n) \leqslant \rho(x_i, y) \leqslant \rho(x, x_i) + \rho(x, y) < 2^{-n+2} r.$$

Thus $i \notin B$, so $i \in A$ and therefore $x_i \in F_{n+1}$. Since $\rho(x, x_i) < 2^{-(n+1)+1}r$, the set F_{n+1} therefore satisfies the appropriate instance of (a). This completes the inductive construction.

Let K be the closure of $\bigcup_{n \geqslant 1} F_n$ in X. We see from (a) that $B(x_0, r) \subset K$. On the other hand, given $x \in K$ and a positive integer n, we can find m and $y \in F_m$ such that $\rho(x, y) < 2^{-n+4}r$. If $m \geqslant n$, then by (b), there exist points $y_m = y$, $y_{m-1} \in F_{m-1}, \ldots, y_n \in F_n$ such that $\rho(y_{i+1}, y_i) < 2^{-i+3}r$ for $n \leqslant i \leqslant m-1$. Thus

$$\rho(y, F_n) \leqslant \rho(y, y_n) \leqslant \sum_{i=n}^{m-1} \rho(y_{i+1}, y_i) < \sum_{i=n}^{\infty} 2^{-i+3}r = 2^{-n+4}r \tag{2.4}$$

and therefore

$$\rho(x, F_n) \leqslant \rho(x, y) + \rho(y, F_n) < 2^{-n+4}r + 2^{-n+4}r = 2^{-n+5}r.$$

It follows that $\bigcup_{i=1}^{n} F_i$ is a finitely enumerable $2^{-n+5}r$-approximation to K. Since n is arbitrary, we conclude that K is totally bounded.

Finally, taking $n = 1$ in (2.4), we see that $\rho(y, x_0) < 8r$ for each y in $\bigcup_{n \geqslant 1} F_n$; whence $K \subset \overline{B}(x_0, 8r)$. □

Corollary 2.2.12. *If X is a totally bounded space, then for each $\varepsilon > 0$ there exist totally bounded sets $K_1, \ldots, K_n$, each of diameter less than or equal to ε, such that $X = \bigcup_{i=1}^{n} K_i$.*

Proof. Given $\varepsilon > 0$, construct an $\varepsilon/16$-approximation $\{x_1, \ldots, x_n\}$ to X. By Proposition 2.2.11, for each i in $\{1, \ldots, n\}$ there exists a closed, totally bounded set K_i such that $B(x_i, \varepsilon/16) \subset K_i \subset \overline{B}(x_i, \varepsilon/2)$. Clearly, $X = \bigcup_{i=1}^{n} K_i$. Also, $\rho(x, y) \leqslant \varepsilon$ for all x, y in K_i, so $\operatorname{diam}(K_i) \leqslant \varepsilon$. □

A property P, applicable to certain elements of a set S, is said to hold *for all but countably many* x in S if there exists a sequence $(x_n)_{n \geqslant 1}$ in S such that $P(x)$ holds whenever $x \in S$ and $x \neq x_n$ for each n. The sequence $(x_n)_{n \geqslant 1}$ is then called the *excluded sequence,* and the elements x such that $x \neq x_n$ for each n are said to be *admissible,* for the property P.

Theorem 2.2.13. *Let f be a uniformly continuous mapping of a totally bounded metric space X into $\mathbb{R}$. Then for all but countably many $r \in \mathbb{R}$ the set*

$$X(f, r) = \{x \in X : f(x) \leqslant r\}$$

is either totally bounded or empty.

Proof. By Corollary 2.2.12, for each positive integer k there exist a positive integer n_k and totally bounded sets $X_{k,j}$ $(1 \leqslant j \leqslant n_k)$, each of diameter less than $1/k$, whose union is X. Define the excluded sequence $(r_n)_{n\geqslant 1}$ to be an enumeration of the real numbers

$$c_{k,j} = \inf \{f(x) : x \in X_{k,j}\} \quad (k \geqslant 1;\ 1 \leqslant j \leqslant n_k).$$

(Note that the numbers $c_{k,j}$ exist in view of Corollary 2.2.7.) Let $r \neq r_n$ for each n. If $c_{k,j} > r$ for $1 \leqslant j \leqslant n_k$, then $X(f,r) = \varnothing$. So we may assume that there exists $\nu \leqslant n_k$ such that $c_{k,j} < r$ for $1 \leqslant j \leqslant \nu$, and $c_{k,j} > r$ for $\nu < j \leqslant n_k$. For each $j \leqslant \nu$ choose $x_{k,j} \in X_{k,j}$ such that $f(x_{k,j}) < r$. For all such k, j and all $x \in X_{k,j}$ we have

$$\rho(x, x_{k,j}) \leqslant \operatorname{diam}(X_{k,j}) < \frac{1}{k}. \tag{2.5}$$

Consider any $x \in X(f,r)$ and any positive integer k. Choose j with $1 \leqslant j \leqslant n_k$ and $x \in X_{k,j}$. Then $c_{k,j} \leqslant f(x) \leqslant r$, so $c_{k,j} < r$ (since $c_{k,j} \neq r$) and therefore $j \leqslant \nu$. By (2.5), $\rho(x, x_{k,j}) < 1/k$. Hence $\{x_{k,j} : 1 \leqslant j \leqslant \nu\}$ is a finitely enumerable $1/k$-approximation to $X(f,r)$. Since k is arbitrary, it follows that $X(f,r)$ is totally bounded. □

A complete, totally bounded metric space X is said to be *compact.* The bounded closed intervals $[a,b]$ in $\mathbb{R}$, and the closed balls in $\mathbb{C}$, are compact; the product of finitely many compact spaces is compact. A compact subset of a metric space is both closed and (by Proposition 2.2.9) located; and a closed, located subset of a compact space is compact (see Proposition 2.2.10).

Corollary 2.2.14. *Under the hypotheses of* Theorem 2.2.13, *if X is compact, then $X(f,r)$ is either compact or empty for all but countably many $r \in \mathbb{R}$.*

Proof. In view of Theorem 2.2.13, it suffices to note that, by the uniform continuity of f, the set $X(f,r)$ is closed in X, and therefore complete, for each admissible r. □

We say that two subsets S, T of a metric space X are *apart* from each other, and we write $S \bowtie T$, if there exists $r > 0$ such that $\rho(s,t) \geqslant r$ for all $s \in S$ and $t \in T$. A mapping $f : X \longrightarrow Y$ between metric spaces is said to be *strongly continuous* if for all subsets S, T of X with $f(S) \bowtie f(T)$ we have $S \bowtie T$. It is simple to show that uniform continuity implies strong continuity. The following partial converse will be used shortly.

Proposition 2.2.15. *Let f be a strongly continuous mapping of a metric space X onto a totally bounded metric space Y. Then f is uniformly continuous.*

Proof. Given $\varepsilon > 0$, construct an $\varepsilon/8$-approximation $\{f(x_1), \dots, f(x_n)\}$ to Y, and define

$$Y_i = B\left(f(x_i), \tfrac{\varepsilon}{8}\right), \quad X_i = f^{-1}(Y_i).$$

Write $\{(i,j) : 1 \leqslant i,j \leqslant n\}$ as the union of subsets P and Q such that

$$(i,j) \in P \Longrightarrow \rho\left(f(x_i), f(x_j)\right) < \frac{3\varepsilon}{4},$$
$$(i,j) \in Q \Longrightarrow \rho\left(f(x_i), f(x_j)\right) > \frac{\varepsilon}{2}.$$

If $(i,j) \in P$, then, by the triangle inequality, $\rho\left(f(x), f(x')\right) < \varepsilon$ for all $x, x' \in X$. So we may assume that there exists $(i,j) \in Q$. For such (i,j) and any element (y, y') of $Y_i \times Y_j$, since $\rho\left(f(x_i), y\right) < \varepsilon/8$ and $\rho\left(y', f(x_j)\right) < \varepsilon/8$, we have

$$\rho(y, y') \geqslant \rho\left(f(x_i), f(x_j)\right) - \rho\left(f(x_i), y\right) - \rho\left(y', f(x_j)\right) > \frac{\varepsilon}{4}.$$

Hence $Y_i \bowtie Y_j$. The strong continuity of f now yields $X_i \bowtie X_j$; whence there exists $r_{i,j} > 0$ such that $\rho\left(x, x'\right) \geqslant r_{i,j}$ for all $x \in X_i$ and $x' \in X_j$. Let

$$\delta = \min\left\{r_{i,j} : (i,j) \in Q\right\} > 0,$$

and consider points x, x' of X with $\rho\left(x, x'\right) < \delta$. Choose i, j such that $f(x) \in Y_i$ and $f(y) \in Y_j$. If $(i,j) \in Q$, then $\rho\left(x, x'\right) \geqslant r_{i,j} \geqslant \delta$, a contradiction. Hence $(i,j) \in P$; so $\rho\left(f(x_i), f(x_j)\right) < 3\varepsilon/4$ and therefore $\rho\left(f(x), f(x')\right) < \varepsilon$. □

A mapping $f : X \longrightarrow Y$ between metric spaces is said to be *strongly injective*, or a *strong injection*, if it is one-one and the inverse mapping is strongly continuous; in other words, if for all subsets S, T of X with $S \bowtie T$ we have $f(S) \bowtie f(T)$. Every strongly injective mapping is injective.

The following is a constructive substitute for the classical theorem that a continuous one-one mapping of a compact metric space into a metric space has compact range and a continuous inverse (a theorem that is false in RUSS).

Proposition 2.2.16. *Let f be a uniformly continuous strong injection of a compact metric space X into a metric space Y. Then f^{-1} is uniformly continuous and strongly injective on $f(X)$, and $f(X)$ is compact.*

Proof. The mapping $f^{-1} : f(X) \longrightarrow X$ is strongly continuous, and X is totally bounded. Hence, by Proposition 2.2.15, f^{-1} is uniformly continuous on $f(X)$. Since f is uniformly continuous and hence strongly continuous, f^{-1} is also strongly injective. By Proposition 2.2.6, $f(X)$ is totally bounded; so it remains to prove that it is complete. Accordingly, let $(f(x_n))_{n\geqslant 1}$ be a Cauchy sequence in $f(X)$. Using the uniform continuity of f^{-1}, we see that $(x_n)_{n\geqslant 1}$ is a Cauchy sequence in X. Since X is complete, $(x_n)_{n\geqslant 1}$ converges to a limit x in X. Finally, since f is continuous, $(f(x_n))_{n\geqslant 1}$ converges to $f(x)$ in $f(X)$. □

Let $\mathcal{K}(X)$ denote the set of all compact subsets of a compact metric space X. For all $A, B \in \mathcal{K}(X)$, since the map $x \rightsquigarrow \rho(x, B)$ is defined (B being located) and is uniformly continuous on A, the number

$$m_{A,B} = \inf \{\rho(x, B) : x \in A\}$$

exists, by Corollary 2.2.7. Likewise, $m_{B,A}$ exists. It is left as an exercise to show that the mapping

$$(A, B) \rightsquigarrow \max \{m_{A,B}, m_{B,A}\}$$

from $\mathcal{K}(X) \times \mathcal{K}(X)$ into the nonnegative real line

$$\mathbb{R}^{0+} = \{x \in \mathbb{R} : x \geqslant 0\}$$

is a metric—we call it the *Hausdorff metric*—with respect to which the space $\mathcal{K}(X)$ is compact.

An inhabited metric space X is said to be

- *locally totally bounded* if each bounded subset of X is contained in a totally bounded subset;
- *locally compact* if it is both locally totally bounded and complete.

Every compact space is locally compact. The spaces $\mathbb{R}$ and $\mathbb{C}$, and the product spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, are locally compact. A metric space X is locally compact if and only if every bounded subset of X is contained in a compact set.

The following lemma prepares the way for our next theorem, which deals with certain fundamental properties of a locally totally bounded space.

Lemma 2.2.17. *Let Y be a located subset of a metric space X, and T a totally bounded subset of X that intersects Y. Then there exists a totally bounded set S such that $T \cap Y \subset S \subset Y$.*

Proof. Using Theorem 2.2.13, construct a sequence $(\alpha_n)_{n \geqslant 1}$ such that for each positive integer n, $0 < \alpha_{n+1} < \alpha_n < 1/n$ and

$$T_n = \{x \in T : \rho(x, Y) \leqslant \alpha_n\}$$

is totally bounded. Let F_n be a finite $1/n$-approximation to T_n, and construct a mapping $\phi_n : F_n \longrightarrow Y$ such that $\rho(x, \phi_n(x)) < 1/n$ for each $x \in F_n$; then set

$$S_n = \{\phi_n(x) : x \in F_n\}.$$

Let S be the closure of $\bigcup_{n \geqslant 1} S_n$ in Y. We show that S_N is a $3/N$-approximation to $\bigcup_{k \geqslant N} S_k$. Let $k \geqslant N$ and $y \in S_k$. Then $y = \phi_k(x) \in Y$ for some $x \in F_k$. Since $F_k \subset T_k \subset T_N$, there exists $z \in F_N$ such that $\rho(x, z) < 1/N$; whence

$$\rho(y, \phi_N(z)) \leqslant \rho(\phi_k(x), x) + \rho(x, z) + \rho(z, \phi_N(z)) < \frac{1}{k} + \frac{1}{N} + \frac{1}{N} \leqslant \frac{3}{N}.$$

We now see that $\bigcup_{k=1}^{N} S_k$ is a $3/N$-approximation to $\bigcup_{n \geqslant 1} S_n$. So $\bigcup_{n \geqslant 1} S_n$, and therefore S, is totally bounded.

If $x \in T \cap Y$, then for each n we have $x \in T_n$; so there exists $z \in F_n$ such that $\rho(x, z) < 1/n$ and therefore $\rho(x, \phi_n(z)) < 2/n$. Thus $\rho(x, S) < 2/n$ for each positive integer n. Since S is closed in Y, it follows that x belongs to S. □

Proposition 2.2.18. *Let Y be an inhabited subset of a metric space X.*

(a) *If Y is locally totally bounded, then it is located.*

(b) *If X is locally totally bounded and Y is located, then Y is locally totally bounded.*

Proof. Assume first that Y is locally totally bounded. Let $y_0 \in Y$ and $x \in X$. The set

$$B = \{y \in Y : \rho(y, y_0) \leqslant 2\rho(x, y_0) + 1\},$$

being bounded in Y, is contained in a totally bounded subset K of Y. By Theorem 2.2.13, there exists $r > 2\rho(x, y_0) + 1$ such that the set

$$T = \{y \in K : \rho(y, y_0) \leqslant r\}$$

is totally bounded and hence located. Then $\rho(x, T) \leqslant \rho(x, y_0) < \rho(x, y_0) + 1$, so for each y in Y, either $\rho(x, y) > \rho(x, T)$ or $\rho(x, y) < \rho(x, y_0) + 1$. In the latter case,

$$\rho(y, y_0) \leqslant \rho(x, y) + \rho(x, y_0) \leqslant 2\rho(x, y_0) + 1 < r,$$

so $y \in T$ and therefore $\rho(x, y) \geqslant \rho(x, T)$. It follows that $\rho(x, Y)$ exists and equals $\rho(x, T)$. Thus Y is located.

Now assume that X is locally totally bounded and that Y is located in X. Let B be a bounded subset of Y. Then there exists a totally bounded subset T of X such that $B \subset T$. By Lemma 2.2.17, there exists a totally bounded subset S of Y such that $T \cap Y \subset S$ and therefore $B \subset S$. Thus Y is locally totally bounded. □

2.3 Normed Linear Spaces

Let $\mathbb{K}$ stand for either $\mathbb{R}$ or $\mathbb{C}$, and let X be a linear space over $\mathbb{K}$. An inequality relation $\neq$ on X is said to be *compatible with the linear structure* on X if for all $x, y \in X$ and all $t \in \mathbb{K}$,

$$\begin{aligned} x \neq y &\Longleftrightarrow x - y \neq 0, \\ x + y \neq 0 &\Longrightarrow x \neq 0 \vee y \neq 0, \text{ and} \\ tx \neq 0 &\Longrightarrow t \neq 0 \wedge x \neq 0. \end{aligned}$$

It readily follows from the first of these properties that

$$x \neq y \Longrightarrow \forall z \in X \, (x + z \neq y + z) \, .$$

The requirement of compatibility between the inequality and the linear structure is a natural one and is automatically fulfilled by the denial inequality under classical logic. Constructively, it is certainly true that the denial inequality fulfils the first of the three requirements for compatibility; but unless we accept Markov's principle we cannot expect to prove that it satisfies the second or the third. In the case $X = \mathbb{K}$ the standard inequality is compatible with the linear structure.

From now on, unless we state otherwise, when we refer to a *linear space* we mean one that is equipped with a compatible inequality. Let X be such a space. By a *seminorm* on X we mean a mapping $\| \ \| : x \rightsquigarrow \|x\|$ of X into $\mathbb{R}^{0+}$ such that for all x, y in X and all t in $\mathbb{K}$,

- $\|x\| > 0 \Longrightarrow x \neq 0$,
- $\|tx\| = |t| \, \|x\|$, and
- $\|x + y\| \leqslant \|x\| + \|y\|$.

We call the pair $(X, \| \ \|)$—or, when no confusion is likely, just X itself—a *seminormed (linear) space* over $\mathbb{K}$. We say that the seminormed space X is *real* or *complex*, depending on whether $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$. If $x \in X$ and $\|x\| > 0$, then x is called a *nonzero vector*; if $\|x\| = 1$, then x is called a *unit vector*. We call the seminormed space X *nontrivial* if it contains a nonzero vector.

If the inequality on the seminormed space X satisfies

$$x \neq 0 \Longleftrightarrow \|x\| > 0, \tag{2.6}$$

then we call $\| \ \|$ a *norm* on X, $\|x\|$ the *norm of the vector* x, and $(X, \| \ \|)$—or just X—a *normed (linear) space* over $\mathbb{K}$. Note that every seminorm $\| \ \|$ on a linear space X induces an inequality relation—namely, the one defined by (2.6)—with respect to which $\| \ \|$ becomes a norm.

Let X be a normed space. Then the mapping $(x, y) \rightsquigarrow \|x - y\|$ of $X \times X$ into $\mathbb{R}$ is a metric on X, and is said to be *associated* with the norm on X. When we consider X as a metric space, it is understood that we are referring to the metric, usually denoted by ρ, associated with the given norm on X. Note that the inequality corresponding to the metric associated with the norm on X is just the original inequality on X.

The *unit ball* of X is the closed ball with centre 0 and radius 1,

$$B_X = \overline{B}_X(0,1) = \overline{B}(0,1) = \{x \in X : \|x\| \leqslant 1\},$$

relative to that metric. This ball, like any open or closed ball in a normed space, is located.

It is a simple consequence of the triangle inequality that

$$|\,\|x\| - \|y\|\,| \leqslant \|x - y\|$$

for all vectors x, y in a normed space X. It follows from this that if a sequence $(x_n)_{n\geqslant 1}$ converges to a limit x in X, then the sequence $(\|x_n\|)_{n\geqslant 1}$ converges to $\|x\|$ in $\mathbb{R}$.

Perhaps the simplest examples of a norm are the following ones on $\mathbb{K}^n$:

$$(x_1, \dots, x_n) \rightsquigarrow \max\{|x_1|, \dots, |x_n|\},$$
$$(x_1, \dots, x_n) \rightsquigarrow \sqrt{|x_1|^2 + \cdots + |x_n|^2},$$
$$(x_1, \dots, x_n) \rightsquigarrow |x_1| + \cdots + |x_n|.$$

The second of these is called the *Euclidean norm* on $\mathbb{K}^n$, which, when equipped with that norm, is called *Euclidean n-space* or simply *Euclidean space.*

If S is a compact metric space, then for each uniformly continuous map $f : S \longrightarrow \mathbb{K}$ the *sup norm*

$$\|f\| = \sup\{|f(x)| : x \in S\}$$

is well defined, by Corollary 2.2.7. It is easy to see that the mapping $f \rightsquigarrow \|f\|$ is a norm on the space $\mathcal{C}(S, \mathbb{K})$ of uniformly continuous functions from S to $\mathbb{K}$, taken with pointwise operations of addition and multiplication-by-scalars. We usually denote $\mathcal{C}(S, \mathbb{C})$ by $\mathcal{C}(S)$. Convergence and Cauchyness with respect to the sup norm on $\mathcal{C}(S, \mathbb{K})$, where S is a compact metric space, are called *uniform convergence* and *uniform Cauchyness* respectively. The space $\mathcal{C}(S, \mathbb{K})$ is an example of a *Banach space*: that is, a complete normed space. The standard proof of this in classical analysis is constructive, and is left as an exercise. The Euclidean space $\mathbb{K}^n$ is also a Banach space.

Let X_1, X_2 be normed spaces over $\mathbb{K}$, and recall that the standard operations of addition and multiplication-by-scalars on the product vector space $X = X_1 \times X_2$ are given by

$$(x_1, x_2) + (x_1', x_2') = (x_1 + x_1',\, x_2 + x_2'),$$
$$t\,(x_1, x_2) = (tx_1, tx_2),$$

where $x_i \in X_i,\ x_i' \in X_i$, and $t \in \mathbb{K}$. It is easy to verify that the mapping

$$(x_1, x_2) \rightsquigarrow \max\{\|x_1\|, \|x_2\|\}$$

is a norm on X, and that the metric associated with this norm is the product metric on X (considered as the product of the metric spaces X_1 and X_2). Taken with this norm, which we call the *product norm*, X is known as the *product of the normed spaces* X_1 and X_2. The product norm and the product space for a finite number of normed spaces are defined analogously.

The proof of the following is left as an exercise.

Proposition 2.3.1. *Let X be a normed space over $\mathbb{K}$. Then*

(a) *the mapping $(x, y) \rightsquigarrow x + y$ is uniformly continuous on $X \times X$;*

(b) *for each $t \in \mathbb{K}$ the mapping $x \rightsquigarrow tx$ is uniformly continuous on X;*

(c) *for each $x \in X$ the mapping $t \rightsquigarrow tx$ is uniformly continuous on $\mathbb{K}$;*

(d) *the mapping $(t, x) \rightsquigarrow tx$ is continuous on $\mathbb{K} \times X$.*

If X is a normed space and S is a linear subset of X, then the restriction to S of the norm on X is a norm on S; taken with this norm, S is called a *normed linear subspace*, or simply a *subspace*, of X. It is left as an exercise to show, from Proposition 2.3.1, that the closure of a subspace in X is also a subspace of X.

Recall that a mapping u between vector spaces X, Y is *linear* if

$$u(x + y) = u(x) + u(y) \quad \text{and} \quad u(tx) = tu(x)$$

whenever $x, y \in X$ and $t \in \mathbb{K}$. For example, the mapping $x \rightsquigarrow Ax$ on $\mathbb{K}^n$, where A is an n-by-n matrix over $\mathbb{K}$, is linear.

If $X = Y$, then we refer to a linear map $u : X \longrightarrow Y$ as an *operator* on X. On the other hand, if $Y = \mathbb{K}$, then u is called a *linear functional* on X.

Continuous linear mappings between normed spaces are the backbone of functional analysis. In order to characterise these maps, we need a simple lemma, whose proof is left to the reader.

Lemma 2.3.2. *For each element x of a normed space X,*

$$\|x\| = \inf \left\{ \frac{1}{|t|} : t \in \mathbb{K}, \ \ t \neq 0, \ \ \|tx\| \leqslant 1 \right\}.$$

Proposition 2.3.3. *The following are equivalent conditions on a linear mapping of a normed space X into a normed space Y:*

(a) *u is continuous at* 0.

(b) *u is continuous on X.*

(c) *u is uniformly continuous on X.*

(d) *$u(B_X)$ is a bounded subset of Y.*

(e) *u maps bounded subsets of X to bounded subsets of Y.*

(f) *There exists a positive number c, called a* **bound** *for u, such that $\|u(x)\| \leqslant c\|x\|$ for all $x \in X$.*

Proof. Suppose that u is continuous at 0. Then there exists $r > 0$ such that

$$\|u(x)\| = \|u(x) - u(0)\| \leqslant 1$$

whenever $\|x\| \leqslant r$. For each nonzero $t \in \mathbb{K}$ with $\|tx\| \leqslant 1$ we have $\|rtx\| \leqslant r$ and therefore

$$\|u(x)\| = \frac{1}{r\,|t|}\,\|u(rtx)\| \leqslant \frac{1}{r\,|t|}.$$

It follows from Lemma 2.3.2 that $\|u(x)\| \leqslant \frac{1}{r}\|x\|$ for all $x \in X$. Hence (a) implies (f).

It is clear that (f) $\Longrightarrow$ (e) $\Longrightarrow$ (d). Next, suppose that there exists $c > 0$ such that $\|u(x)\| \leqslant c$ whenever $\|x\| \leqslant 1$. Given x in X and $\varepsilon > 0$, we have either $x \neq 0$ or $\|x\| < \varepsilon$. In the first case,

$$\|u(x)\| = \|x\| \left\| u\left(\frac{1}{\|x\|}x\right) \right\| \leqslant c\,\|x\|\,.$$

In the second, $\left\|u\left(\varepsilon^{-1}x\right)\right\| \leqslant c$, so

$$\|u(x)\| \leqslant c\varepsilon \leqslant c\,(\|x\| + \varepsilon)\,.$$

Thus $\|u(x)\| \leqslant c\,(\|x\| + \varepsilon)$ in each case. Letting $\varepsilon \longrightarrow 0$, we obtain $\|u(x)\| \leqslant c\,\|x\|$. Thus (f) holds. We now have

$$\|u(x - y)\| \leqslant c\,\|x - y\| \quad (x, y \in X)\,,$$

from which it follows that u is uniformly continuous on X. Thus (d) $\Longrightarrow$ (f) $\Longrightarrow$ (c).

Finally, it is immediate that (c) $\Longrightarrow$ (b) $\Longrightarrow$ (a). □

In view of property (e) of Proposition 2.3.3, we refer to a continuous linear mapping between normed spaces X, Y as a *bounded linear mapping*. In the cases $Y = X$ and $Y = \mathbb{K}$, we use the terms *bounded operator* and *bounded linear functional*, respectively.

Bounded linear functionals are associated with linear subsets that, in the case of a general normed space, correspond to planes in three-dimensional geometry.

A linear subset H of a normed space X is called a *hyperplane* if there exist an *associated vector* $x_0 \in X$ and a positive number c such that

- ▷ $\|x - x_0\| \geqslant c$ for each $x \in H$, and
- ▷ each $x \in X$ is represented (perforce uniquely) in the form $x = tx_0 + y$ with $t \in \mathbb{K}$ and $y \in H$.

Recall that the *kernel* of a linear mapping u between linear spaces X, Y is the set

$$\ker(u) = u^{-1}(0).$$

We say that u is *nonzero* if there exists $x \in X$ such that $u(x) \neq 0$. If u is a nonzero bounded linear functional, then $\ker(u)$ is a hyperplane. To see this, choose $x_0 \in X$ such that $u(x_0) = 1$, and then (by Proposition 2.3.3) $c > 0$ such that $c\,|u(x)| \leqslant \|x\|$ for each $x \in X$. If $x \in \ker(u)$, then

$$\|x - x_0\| \geqslant c\,|u(x - x_0)| = c\,|u(x) - u(x_0)| = c.$$

On the other hand, for each $x \in X$ we have

$$x = u(x)x_0 + (x - u(x)x_0),$$

where $u(x) \in \mathbb{K}$ and $x - u(x)x_0 \in \ker(u)$.

Proposition 2.3.4. *Let X be a normed space, and H a hyperplane in X with associated vector x_0. Then there exists a unique bounded linear functional u on X such that* $\ker(u) = H$ *and* $u(x_0) = 1$.

Proof. Compute $c > 0$ such that $\|x - x_0\| \geqslant c$ for each $x \in H$. For each $x \in X$ there exist unique $u(x) \in \mathbb{K}$ and $\phi(x) \in H$ such that $x = u(x)x_0 + \phi(x)$. By the uniqueness, $u(x_0) = 1$ and $H = \ker(u)$. Moreover, for each $\varepsilon > 0$ we have either

$$|u(x)| \leqslant \frac{1}{c}\,\|x\| + \varepsilon \tag{2.7}$$

or else $u(x) \neq 0$. In the latter case we have

$$\|x\| = |u(x)| \left\| x_0 - \left(-\frac{1}{u(x)}\phi(x) \right) \right\| \geqslant |u(x)|\,c,$$

from which we see that (2.7) holds. Since $\varepsilon > 0$ is arbitrary, we conclude that $1/c$ is a bound for u. □

If $u : X \longrightarrow Y$ is a bounded linear mapping between normed spaces and

$$\|u\| = \sup\{\|u(x)\| : x \in X,\ \|x\| \leqslant 1\} \tag{2.8}$$

exists, we call this number the *norm* of u and we say that u is *normed* or *normable.* In that case,

$$\|u(x)\| \leqslant \|u\| \, \|x\| \quad (x \in X),$$

and if X is nontrivial, then

$$\|u\| = \sup \{\|u(x)\| : x \in X, \ \|x\| = 1\}.$$

Although there is no general criterion for the existence of the norm of a bounded linear mapping between normed spaces, such a criterion—and a very useful one at that—exists for bounded linear functionals. Our proof depends on a preliminary lemma.

Lemma 2.3.5. *Let u be a bounded linear map of a normed space X into a normed space Y. Then* $\ker(u)$ *is located if and only if*

$$s_x = \inf \{t : t > 0, \ u(x) \in tu(B_X)\} \tag{2.9}$$

exists for each $x \in X$. Moreover, if $u(x) \neq 0$, then $s_x > 0$.

Proof. For each $x \in X$ we have

$$\begin{aligned} &\{t > 0 : \exists y \in \ker(u)(\|x - y\| < t)\} \\ &\qquad = \{t > 0 : \exists z \in X \ \ (\|z\| < 1 \text{ and } u\,(x - tz) = 0)\} \\ &\qquad = \{t > 0 : u(x) \in tu(B_X)\}. \end{aligned}$$

The first part of the lemma now follows, since $\rho\,(x, \ker(u))$, if it exists, equals the infimum of the first set, whereas s_x, if it exists, equals the infimum of the third.

Now let $u(x) \neq 0$, and choose $r > 0$ such that if $y \in X$ and $\|y\| < r$, then $|u(y)| < |u(x)|$. Supposing that $s_x < r$, we can find a positive $t < r$ and an element z of B_X such that $u(x) = tu(z) = u(tz)$; but $\|tz\| < r$, so $|u(tz)| < |u(x)|$, a contradiction. Hence $s_x \geqslant r$. □

Proposition 2.3.6. *A nonzero bounded linear functional on a normed space is normed if and only if its kernel is located.*

Proof. Let u be a nonzero bounded linear functional on the normed space X. Supposing first that u is normed—in which case $\|u\| > 0$—consider any $a \in X$. For each $y \in \ker(u)$ we have

$$\|a - y\| \geqslant \frac{|u(a - y)|}{\|u\|} = \frac{|u(a)|}{\|u\|}.$$

On the other hand, if $0 < \varepsilon < \|u\|$ and we choose a unit vector $x \in X$ such that $u(x) > \|u\| - \varepsilon$, then

$$z = a - \frac{u(a)}{u(x)} x \in \ker(u)$$

and

$$\|a - z\| = \frac{|u(a)|}{u(x)} < \frac{|u(a)|}{\|u\| - \varepsilon}.$$

Hence $\rho(a, \ker(u))$ exists and equals $|u(a)| / \|u\|$. Since $a \in X$ is arbitrary, $\ker(u)$ is located.

Conversely, if $\ker(u)$ is located, then by Lemma 2.3.5,

$$s = \inf\{t > 0 : 1 \in tu(B_X)\}$$

exists and is positive. We show that $\|u\| = 1/s$. For each $x \in B_X$ we have either $u(x) \leqslant 1/s$ or $u(x) \neq 0$; in the latter case,

$$\left\| \frac{|u(x)|}{u(x)} x \right\| \leqslant 1$$

and

$$1 = \frac{1}{|u(x)|} u\left(\frac{|u(x)|}{u(x)} x\right),$$

so $1/|u(x)| \geqslant s$ and therefore $|u(x)| \leqslant 1/s$. Thus $1/s$ is a bound for u. On the other hand, given ε with $0 < \varepsilon < 1/s$, choose $t < s/(1 - \varepsilon s)$ and $x \in B_X$ such that $1 = tu(x)$. Then $u(x) = 1/t > 1/s - \varepsilon$. It now follows that $\|u\|$ exists and equals $1/s$. □

We cannot expect the norm of a normed linear functional u to be attained at some vector in the unit ball of the domain space. For a certain important class of spaces, though, attainment of the norm does occur.

A normed space X is said to be *uniformly convex* if for each $\varepsilon > 0$ there exists δ with $0 < \delta < 1$ such that if x, y are unit vectors in X with $\left\|\frac{1}{2}(x + y)\right\| > 1 - \delta$, then $\|x - y\| < \varepsilon$. Every linear subspace of a uniformly convex normed space is itself uniformly convex.

Proposition 2.3.7. *If u is a nonzero normed linear functional on a uniformly convex Banach space X, then there exists a unique unit vector $x \in X$ such that $u(x) = \|u\|$.*

Proof. Construct a sequence $(x_n)_{n \geqslant 1}$ of unit vectors in X such that $u(x_n) \longrightarrow \|u\|$ as $n \longrightarrow \infty$. Given $\varepsilon > 0$, choose $\delta > 0$ as in the definition of "uniformly convex". Since $1 - \frac{\delta}{2} < 1$, there exists a positive integer N such that $\left(1 - \frac{\delta}{2}\right)\|u\| < u(x_n)$ for all $n \geqslant N$. For $m, n \geqslant N$ we then have

$$\begin{aligned}
\|u\| \left\| \frac{1}{2}(x_m + x_n) \right\| &\geqslant \frac{1}{2} |u(x_m + x_n)| \\
&\geqslant u(x_m) - \frac{1}{2} |u(x_m - x_n)| \\
&\geqslant u(x_m) - \frac{1}{2}(\|u\| - u(x_m)) - \frac{1}{2}(\|u\| - u(x_n)) \\
&> \left(1 - \frac{\delta}{2}\right)\|u\| - \frac{\delta}{4}\|u\| - \frac{\delta}{4}\|u\| \\
&= (1-\delta)\|u\| .
\end{aligned}$$

Hence $\left\| \frac{1}{2}(x_m + x_n) \right\| > 1-\delta$ and therefore $\|x_m - x_n\| < \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we conclude that $(x_n)_{n \geqslant 1}$ is a Cauchy sequence. It therefore converges to a limit x in the complete space X. By the continuity of the norm and of u, we have $\|x\| = 1$ and $u(x) = \|u\|$.

Now let y be any unit vector such that $u(y) = \|u\| = u(x)$. For each positive integer n define $z_{2n-1} = x$ and $z_{2n} = y$. Then $(z_n)_{n \geqslant 1}$ is a sequence of unit vectors such that $u(z_n) \longrightarrow \|u\|$; so, by the first part of the proof, $(z_n)_{n \geqslant 1}$ converges to a unit vector, which clearly must equal both x and y. Hence $x = y$. □

We next present an important construction of new normed spaces from old. Let Y be a located subspace of a normed space X, and define new equality and inequality relations on X by

$$\begin{aligned}
x =_{X/Y} x' &\Longleftrightarrow \rho(x - x', Y) = 0, \\
x \neq_{X/Y} x' &\Longleftrightarrow \rho(x - x', Y) > 0.
\end{aligned}$$

Taken with this equality relation, X becomes the *quotient space* of X by Y, and is usually redesignated X/Y. The identity (linear) mapping $x \rightsquigarrow x$ from the original normed space X into X/Y is then called the *canonical injection* $\mathfrak{i}_{X/Y}$ *of* X onto X/Y. If Y is also closed in X, then

$$\|x\|_{X/Y} = \rho(x, Y)$$

defines a norm—the *quotient norm*—on X/Y. Since $\|x\|_{X/Y} \leqslant \|x - 0\| = \|x\|$, the canonical injection $\mathfrak{i}_{X/Y}$ is a bounded linear mapping.

Note that the locatedness of Y is essential for the definition of the equality, inequality, and norm on X/Y.

Proposition 2.3.8. *If Y is a closed, located subspace of a Banach space X, then the quotient space X/Y is a Banach space.*

Proof. Given a Cauchy sequence $(x_n)_{n \geqslant 1}$ in X/Y, choose a strictly increasing sequence $(n_k)_{k \geqslant 1}$ of positive integers such that

$$\left\|x_{n_{k+1}} - x_{n_k}\right\|_{X/Y} < 2^{-k} \quad (k \geqslant 1).$$

Setting $y_1 = 0$, we construct inductively a sequence $(y_k)_{k\geqslant 1}$ in Y such that for each k,

$$\left\|(x_{n_{k+1}} - y_{k+1}) - (x_{n_k} - y_k)\right\| < 2^{-k}. \tag{2.10}$$

Indeed, having constructed elements $y_1, \dots, y_k$ of Y with the applicable properties, we have

$$\inf\left\{\left\|x_{n_{k+1}} - (x_{n_k} - y_k) - y\right\| : y \in Y\right\} = \left\|x_{n_{k+1}} - x_{n_k}\right\|_{X/Y} < 2^{-k},$$

so there exists $y_{k+1} \in Y$ such that (2.10) holds. We now see from (2.10) that $(x_{n_k} - y_k)_{k\geqslant 1}$ is a Cauchy sequence in the Banach space X; whence it converges to a limit z in X. We then have

$$\|x_{n_k} - z\|_{X/Y} = \|x_{n_k} - z - y_k\|_{X/Y} \leqslant \|(x_{n_k} - y_k) - z\| \longrightarrow 0 \text{ as } k \longrightarrow \infty.$$

Thus the Cauchy sequence $(x_n)_{n\geqslant 1}$ in X/Y has a subsequence that converges in X/Y. It follows that the sequence $(x_n)_{n\geqslant 1}$ itself converges in X/Y. □

Normed spaces form the natural abstract context for infinite series. Given a sequence $(x_n)_{n\geqslant 1}$ of elements of a normed space X, we define the corresponding *series* $\sum_{n=1}^{\infty} x_n$ to be the sequence $(s_n)_{n\geqslant 1}$, where $s_n = \sum_{k=1}^{n} x_k$ is the nth *partial sum* of the series. The series $\sum_{n=1}^{\infty} x_n$ is said to be

- ▷ *convergent* if the sequence $(s_n)_{n\geqslant 1}$ converges to a limit s in X, called the *sum* of the series;
- ▷ *absolutely convergent* if the series $\sum_{n=1}^{\infty} \|x_n\|$ is convergent in $\mathbb{R}$.

In the first case we write $\sum_{n=1}^{\infty} x_n = s$.

If $(a_n)_{n\geqslant 1}$ and $(b_n)_{n\geqslant 1}$ are sequences of real numbers such that $0 < a_n \leqslant b_n$ for each n, and if $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges (the *comparison test*). For, given $\varepsilon > 0$, since the partial sums of a convergent series form a Cauchy sequence we can find N such that

$$0 < \sum_{n=j}^{k} b_n < \varepsilon$$

whenever $k > j \geqslant N$. For all such j, k we then have $0 < \sum_{n=j}^{k} a_j < \varepsilon$. So the partial sums of $\sum_{n=1}^{\infty} a_n$ form a Cauchy sequence, which converges by the completeness of $\mathbb{R}$.

A particular case of this occurs when $b_n = r^n$ for some fixed r with $|r| < 1$: in that case, $\sum_{n=1}^{\infty} b_n$ is a geometric series, and $\sum_{n=1}^{\infty} a_n$ converges to a sum at most $\frac{r}{1-r}$.

We shall resume consideration of normed spaces in Chapter 4.

Exercises

1. Prove that two real numbers x and y are equal if and only if $x \cup y$ is a real number. (This is not a typographical error: remember, a real number is actually a set.)

2. Prove the following for real numbers x, y, z:

$$\begin{aligned} x > y &\Longrightarrow x + z > y + z, \\ (x > 0 \wedge y > 0) &\Longrightarrow xy > 0, \\ xy \neq 0 &\Longrightarrow (x \neq 0 \wedge y \neq 0) . \end{aligned}$$

3. Prove that for all $x, y \in \mathbb{R}$ the sets $\max\{x, y\}$ and $\min\{x, y\}$ are real numbers; that $\min\{x, y\} \leqslant x \leqslant \max\{x, y\}$; that $x = \min\{x, y\}$ if and only if $x \leqslant y$; and that $x = \max\{x, y\}$ if and only if $x \geqslant y$. Prove also that $\max\{x, y\} < \varepsilon$ if and only if $x < \varepsilon$ and $y < \varepsilon$; that $\max\{x, y\} \leqslant \varepsilon$ if and only if $x \leqslant \varepsilon$ and $y \leqslant \varepsilon$; and that $\max\{x, y\} > z$ if and only if either $x > z$ or $y > z$.

4. Prove that $|x| = x$ if and only if $x \geqslant 0$, and that $|x| = -x$ if and only if $x \leqslant 0$. Prove also that $|x| < y$ if and only if $-y < x < y$, and that $|x| \leqslant y$ if and only if $-y \leqslant x \leqslant y$.

5. Prove *Archimedes' axiom* in the following form: If $x > 0$ and $y \geqslant 0$, then there exists an integer n such that $nx > y$.

6. Let x, y be real numbers with $y \neq 0$. Prove that x/y is a real number and that $y/y = 1$.

7. Prove that the statement

$$\forall x \in \mathbb{R}\, (\neg\, (x = 0) \Longrightarrow \exists y \in \mathbb{R}\, (xy = 1))$$

implies Markov's principle.

8. Prove that the mapping $(x, y) \rightsquigarrow x + y$ is a function on $\mathbb{R} \times \mathbb{R}$, and that it is strongly extensional. Prove analogous results for the mappings $x \rightsquigarrow -x$, $(x, y) \rightsquigarrow xy$, $(x, y) \rightsquigarrow x/y$ (for $y \neq 0$), $x \rightsquigarrow |x|$, $(x, y) \rightsquigarrow \min\{x, y\}$, and $(x, y) \rightsquigarrow \max\{x, y\}$.

9. Give Brouwerian examples to show that each of the following statements is essentially nonconstructive:

   (a) If $S \subset \mathbb{R}$ is closed, then $\sim S$ is open in $\mathbb{R}$.

   (b) If $S \subset \mathbb{R}$ and $\sim S$ is open, then S is closed in $\mathbb{R}$.

10. Let X be a separable metric space. Prove that (i) dense subsets and (ii) located subsets of X are separable.

11. Show that the statement "the metric complement of every located subset of $\mathbb{R}$ is located" implies the *weak law of excluded middle (WLEM)*, $\neg P \vee \neg\neg P$.

12. Prove that the closure of a totally bounded subset of a metric space X is totally bounded, and that if a subset S of X contains a dense totally bounded set, then S itself is totally bounded.

13. Prove Corollary 2.2.8 directly, without using Proposition 2.2.5, Proposition 2.2.6, or Corollary 2.2.7.

14. Let h be a mapping of a metric space into a totally bounded space X such that $f \circ h$ is uniformly continuous for each uniformly continuous map $f : X \longrightarrow \mathbb{R}$. Prove that h is uniformly continuous.

15. Let h be a mapping of a compact metric space X into a metric space Y such that $f \circ h$ is uniformly continuous on X for each uniformly continuous mapping $f : Y \longrightarrow \mathbb{R}$. Prove that h is uniformly continuous if and only if its range is totally bounded. Prove also that if Y is locally compact, then h is uniformly continuous.

16. Let $a < b$, and let S, T be inhabited open subsets of $\mathbb{R}$ such that $[a, b] \subset S \cup T$. Prove that $S \cap T$ is inhabited.

17. Let $f : [a, b] \longrightarrow \mathbb{R}$ be sequentially continuous, with $f(a) \leqslant f(b)$. Prove that the range of f is dense in $[f(a), f(b)]$.

18. Let I be a bounded interval, and $f : I \longrightarrow \mathbb{R}$ an increasing sequentially continuous function. Prove that f is uniformly continuous on I.

19. Let S be a dense subset of a metric space X, and f a uniformly continuous mapping of S into a complete metric space Y. Prove that f has a uniformly continuous extension to a mapping of X into Y; that is, a uniformly continuous mapping $F : X \longrightarrow Y$ such that $F(x) = f(x)$ for all $x \in S$.

20. Prove that the Hausdorff metric on the set $\mathcal{K}(X)$ of compact subsets of a compact metric space X is indeed a metric, and that $\mathcal{K}(X)$ is complete with respect to it. Why do we need the sets to be compact in order to be sure that ρ defines a metric?

21. Let X be a compact metric space, $\mathcal{K}(X)$ the compact metric space of all compact subsets of X (see the preceding exercise), and $\mathcal{E}$ the set of all pairs (x, K) where $K \in \mathcal{K}(X)$ and $x \in K$. Define a metric d on $\mathcal{E}$ by
$$d\left(\left(x, K\right), \left(x', K'\right)\right) = \rho\left(x, x'\right) + \rho\left(K, K'\right),$$
where the second ρ on the right-hand side denotes the Hausdorff metric on $\mathcal{K}(X)$. Prove that $\mathcal{E}$ is compact with respect to this metric d.

22. Let X be a metric space, and $(K_n)_{n\geqslant 1}$ a decreasing sequence of compact subsets of X whose diameters converge to 0. Prove that $\bigcap_{n\geqslant 1} K_n$ is compact. Does the conclusion hold without the hypothesis that the diameters converge to 0?

23. Under the hypotheses of Corollary 2.2.14, prove that if $r \in \mathbb{R}$ is admissible and $\varepsilon > 0$, then there exists $\delta > 0$ such that for each admissible r' with $|r - r'| < \delta$,
$$\rho\left(X(f, r), X(f, r')\right) < \varepsilon,$$
where ρ is the Hausdorff metric on $\mathcal{K}(X)$.

24. Give a Brouwerian example to show that we cannot drop the condition "for all but countably many" from the conclusion of Theorem 2.2.13.

25. Prove that the product of finitely many (locally) compact metric spaces is (locally) compact.

26. Let f be a mapping from a metric space (X, ρ) into $\mathbb{R}$ with the following properties:

    (a) the set $X(f, r)$ is compact for certain arbitrarily large real numbers r;

    (b) f is uniformly continuous on each of the sets $X(f, r)$.

    Define a new metric ρ_0 on X by
$$\rho_0\left(x, x'\right) = \rho\left(x, x'\right) + \left|f\left(x\right) - f\left(x'\right)\right| \quad \left(x, x' \in X\right).$$
    Prove that (X, ρ_0) is a locally compact metric space.

27. Let S be a located subset of a locally compact metric space (Y, ρ) such that $X = -S$ is inhabited. Let the mapping $h : Y \longrightarrow \mathbb{R}$ be uniformly continuous on each bounded subset of Y, and such that the set $Y(h, r)$ is bounded for each $r \in \mathbb{R}$. Prove that the mapping $f : x \rightsquigarrow h\left(x\right) + \rho\left(x, S\right)^{-1}$ on X satisfies conditions (a) and (b) of the preceding exercise.

28. Let X be a uniformly convex Banach space, and u a normed linear functional on X. Prove that for any two distinct unit vectors x, y in X either $|u(x)| < \|u\|$ or $|u(y)| < \|u\|$. (This is a strong form of the uniqueness result in Proposition 2.3.7.)

29. Prove Proposition 2.3.1.

30. Prove Lemma 2.3.2.

Notes

In view of our earlier remarks on the status of the power set, it may seem strange, if not perverse, to define a real number as a set. We do not believe that our definition will give rise to any logical problems, since in practice a real number is likely to be given explicitly by interval approximations, rather than by any appeal to the existence of the full power set of $\mathbb{Q} \times \mathbb{Q}$. In fact, many, if not most, real numbers will actually arise as limits of sequences of rational approximations, from which it is straightforward to construct a set of the type required by our definition of "real number".

Bishop defined a real number to be a sequence $(x_n)_{n\geqslant 1}$ of rational numbers that is *regular* in the sense that $|x_m - x_n| < 1/m + 1/n$ for all m and n, two such sequences $(x_n)_{n\geqslant 1}$ and $(y_n)_{n\geqslant 1}$ being called equal if $|x_n - y_n| < 2/n$ for each n. For an axiomatic development of $\mathbb{R}$ see [19].

For us, $\neq$ normally has a stronger meaning than the denial of equality. Some authors, notably the intuitionists, use # instead of $\neq$ to denote an inequality relation, and $\neq$ to denote the denial inequality.

Bishop originally used the word "continuous" to describe mappings that are uniformly continuous on compact sets. While this gets round the independence of the uniform continuity theorem relative to BISH, it has the disadvantage that we cannot prove that the composition of continuous functions is continuous, since the image of a compact set, although totally bounded, cannot generally be proved complete. (To be more precise, in RUSS there is a uniformly continuous mapping of $[0,1]$ onto $(0,1]$.) In a later, unpublished manuscript, *The neat category of stratified spaces*, Bishop introduced the notion of a *compact image:* a subset of a metric space that is the image of a compact metric space under a uniformly continuous mapping. He then defined a mapping $f : X \longrightarrow Y$ between metric spaces to be continuous if it is uniformly continuous *near each compact image*, in the following sense: for each compact image $S \subset X$ and each $\varepsilon > 0$ there exists $\delta > 0$ such that $\rho(f(x), f(x')) < \varepsilon$ whenever $x \in X, x' \in S$, and $\rho(x, x') < \delta$. Compositions of functions that are continuous in this sense are continuous; but the interval $(0,1]$ is a compact image in RUSS, so we cannot prove that the mapping $x \rightsquigarrow 1/x$ on $(0,1]$ is uniformly continuous near compact images. It seems that without Brouwer's fan theorem (see Chapter 5 of [34]), it is impossible to come up with a definition of "continuous" in BISH, other than the usual one of "pointwise continuous", that will satisfy all the conditions that one might wish for. For more on this, see [81] and [89].

In this connection, Bishop's statement that "The concept of a pointwise continuous function is not relevant" ([9], pages ix–x) no longer seems appropriate: there are instances—for example, in the theory of operators between Banach spaces—where sequential convergence is both useful and the best we can hope to prove within BISH; see Chapter 6.

Some authors remove the word "inhabited" from the definitions of "located" and "totally bounded", thereby allowing the empty set to be both located and totally bounded. For us, located sets and totally bounded sets are inhabited, by definition.

Why do we define "compact" as we have done? Classically, compactness is defined in terms of the Heine–Borel–Lebesgue covering property: every open cover contains a finite subcover. In Brouwer's intuitionistic mathematics, as classically, the interval $[0, 1]$ has this open-cover compactness property; but Brouwer's proof depends on his fan theorem, the addition of which to intuitionistic logic would lose us some of the flexibility of interpretation of our constructive mathematics. On the other hand, if we add the Church–Markov–Turing thesis to intuitionistic logic, then we can prove that $[0, 1]$ does not have the Heine–Borel–Lebesgue covering property ([34], page 60, Theorem (4.1)). So that property holds for $[0, 1]$ in one model of our constructive mathematics but fails to hold in another. Therefore, without adding some principle to BISH, we cannot prove or disprove that it holds for $[0, 1]$.

A second classical compactness property, equivalent to the Heine–Borel–Lebesgue one, is that of sequential compactness: every sequence in the space has a convergent subsequence. This fails even for the pair set $\{0, 1\}$ in the constructive setting. Thus, in looking for a workable constructive notion of compactness for metric spaces, we are reduced to that of completeness plus total boundedness, which is classically equivalent to the other two considered above. Fortunately, completeness and total boundedness together make a combination sufficiently powerful for most constructive purposes.

Proposition 2.2.11 appears in *Aspects of Constructivism*, the unpublished notes of colloquium lectures given by Bishop at New Mexico State University in December 1972. It provides a neat proof of Corollary 2.2.12. The proof of Lemma 2.2.17 is new, although not unlike that of Lemma (4.10) on page 33 of [34].

Exercise 14 provides a constructive version of the classical theorem that the unique uniform structure $\mathcal{U}$ compatible with the given topology on a compact Hausdorff space X is induced by the $\mathcal{U}$-uniformly continuous mappings of X into $\mathbb{R}$. Note, however, that we require X to be only totally bounded, not compact.

The notion of strong continuity is best studied in the context of an apartness space: that is, an inhabited set X with an inequality and a binary relation $\bowtie$ of apartness between sets that satisfies certain natural axioms. For more on apartness spaces see [38, 40].

Exercise 16 describes one version of the connectedness of the interval $[a, b]$. For other, constructively inequivalent, types of connectedness see [18].

The result in Exercise 17 is perhaps the most general constructive intermediate value theorem.

A linear map $u : X \longrightarrow Y$ between normed spaces is said to be *compact* if $u(B_X)$ is a totally bounded subset of Y. Proposition 2.3.6 can be generalised as follows: A bounded linear mapping of a normed linear space onto a finite-dimensional Banach space is compact if and only if its kernel is located ([34], page 36, Theorem (5.4)).

Bishop required that a Banach space be separable. We prefer not to depend on separability unless it is absolutely necessary.

We use the term "normed", rather than the usual "normable", for those linear mappings for which the norm exists.

3

The $\boldsymbol{\lambda}$-Technique

Here is more matter for a hot brain.

—Shakespeare, *The Winter's Tale*, act 5, scene 2

In this chapter we discuss a peculiarly constructive technique that, normally under the hypothesis that one or more of the metric spaces under consideration is complete, enables us to prove results that otherwise would require some nonconstructive principle such as LPO, LLPO, or Markov's principle.

3.1 Introduction to the Technique

We begin with another elementary classical result that does not hold in constructive mathematics:

> *For each complex number z there exists $\theta \in [0, 2\pi)$ such that $z = |z|\,\mathrm{e}^{\mathrm{i}\theta}$, and such that if $\theta \neq 0$, then $z \neq 0$.*

We show that this proposition entails LPO. To do so, we consider an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ with at most one term equal to 1. Define a sequence $(z_n)_{n \geqslant 1}$ of complex numbers such that

$$\lambda_n = 0 \Longrightarrow z_n = 0,$$
$$\lambda_n = 1 - \lambda_{n-1} \Longrightarrow z_k = \frac{1}{n}\mathrm{e}^{\mathrm{i}\pi/2} \text{ for all } k \geqslant n.$$

Then $|z_m - z_n| < 1/n$ whenever $m > n$, so $(z_n)_{n \geqslant 1}$ is a Cauchy sequence and therefore converges to a limit z in the complete metric space $\mathbb{C}$. Assume that $z = |z|\,\mathrm{e}^{\mathrm{i}\theta}$ for some $\theta \in [0, 2\pi)$, and that if $\theta \neq 0$, then $z \neq 0$. Either $\theta < \pi/2$ or $\theta > 0$. In the first case we have $\lambda_n = 0$ for all n: for if we suppose that there exists n such that $\lambda_n = 1 - \lambda_{n-1}$, then $z = e^{i\pi/2}/n$ and therefore $\theta = \pi/2$, a contradiction. In the case $\theta > 0$ we have $z \neq 0$, so there exists N such that $z_N \neq 0$; then $\lambda_n = 1$ for some $n \leqslant N$.

The constructive problem occurs when z is near 0 but we cannot decide whether $z = 0$ or $z \neq 0$. Indeed, if $z = 0$ or $z \neq 0$, then we *can* find $\theta \in [0, 2\pi)$ such that $z = |z|\, e^{i\theta}$.

Now, it might be thought that the failure of the modulus–argument decomposition of a general complex number would mean that there was no constructive proof of the existence of square roots in $\mathbb{C}$; for in order to find a square root of z, we normally would write $z = |z|\, e^{i\theta}$ and then compute $\pm\sqrt{|z|}e^{i\theta/2}$. Although this method of finding a square root certainly will not work unless we already can decide that $z = 0$ or $z \neq 0$—which, in general, we cannot—there is a constructive proof of the existence of square roots, one that uses the completeness of $\mathbb{C}$.

To see this, consider any complex number z, and note that for each positive integer n we have either $|z| < 1/n^2$ or $|z| > 1/(n+1)^2$. Thus we can successively construct the terms of an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow |z| < \frac{1}{n^2},$$

$$\lambda_n = 1 \Longrightarrow |z| > \frac{1}{(n+1)^2}.$$

We may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, set $\zeta_n = 0$; if $\lambda_n = 1 - \lambda_{n-1}$, choose $\theta \in [0, 2\pi)$ such that $z = |z|\, e^{i\theta}$, and set $\zeta_k = \sqrt{|z|}e^{i\theta/2}$ for all $k \geqslant n$. Then $(\zeta_n)_{n \geqslant 1}$ is a Cauchy sequence. To see this, let $m \geqslant n$. If $\lambda_m = 0$ or $\lambda_n = 1$, then $\zeta_m = \zeta_n$. If $\lambda_m = 1$ and $\lambda_n = 0$, then there exists a unique k such that $n < k \leqslant m$ and $\lambda_k = 1 - \lambda_{k-1}$; whence $|z| < 1/\left(k-1\right)^2$ and

$$|\zeta_m - \zeta_n| = |\zeta_k - 0| = \sqrt{|z|} < \frac{1}{k-1} \leqslant \frac{1}{n}.$$

Thus, in all cases, $|\zeta_m - \zeta_n| < 1/n$. Also, $\left|\zeta_n^2 - z\right| < 1/n^2$ for each n. Since $\mathbb{C}$ is complete, $(\zeta_n)_{n \geqslant 1}$ converges to a limit $\zeta \in \mathbb{C}$ that clearly satisfies $\zeta^2 = z$.

This technique of "flagging" alternatives by the terms of a binary sequence $(\lambda_n)_{n \geqslant 1}$, constructing an appropriate Cauchy sequence, and using the limit of that sequence to circumvent our inability to make the sort of decision embodied in LPO, LLPO, or Markov's principle, is quite common in constructive mathematics. We present in this chapter a number of results whose proofs are based on this λ-*technique*. Our first such result is *Bishop's lemma*, in whose original proof in [9] (page 177, Lemma 7) the λ-technique first appeared.

Proposition 3.1.1. *Let S be a complete, located subset of a metric space X, and let $x \in X$. There exists $s \in S$ such that if $x \neq s$, then $\rho(x, S) > 0$.*

Proof. Let $s_0 \in S$. If $\rho(x, S) > 1/2$, then we can take $s = s_0$. Hence we may assume that $\rho(x, S) < 1$. Now construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that $\lambda_1 = 0$ and

$$\lambda_n = 0 \Longrightarrow \rho(x, S) < \frac{1}{n},$$
$$\lambda_n = 1 \Longrightarrow \rho(x, S) > \frac{1}{n+1}.$$

If $\lambda_n = 0$, pick $s_n \in S$ such that $\rho(x, s_n) < 1/n$. If $\lambda_n = 1$, set $s_n = s_{n-1}$. Then $\rho(s_m, s_n) < 2/n$ whenever $m \geqslant n$, so $(s_n)_{n \geqslant 1}$ is a Cauchy sequence in S. Since S is complete, this Cauchy sequence converges to a limit $s \in S$ satisfying

$$\rho(s, s_n) \leqslant \frac{2}{n} \qquad (n \geqslant 1).$$

If $x \neq s$, then we can compute a positive integer N such that $\rho(x, s) > 3/N$. If $\lambda_N = 0$, then $\rho(x, s_N) < 1/N$ and so

$$\rho(s, s_N) \geqslant \rho(x, s) - \rho(x, s_N) > \frac{2}{N},$$

a contradiction. Thus $\lambda_N = 1$ and therefore $\rho(x, S) > 0$. □

Bishop's lemma is simple to prove using classical logic. For if $x \in S$, then by taking $s = x$ we render the hypothesis of the implication "if $x \neq s$, then $\rho(x, S) > 0$" false; whereas if $x \notin S$, then the conclusion of that implication holds since S is closed in X, so we may take $s = s_0$. This argument depends on the full law of excluded middle, but in fact only LPO is needed to establish Bishop's lemma classically even when S is merely closed and not necessarily complete. To see this, first construct the sequence $(\lambda_n)_{n \geqslant 1}$ as in our constructive proof of the lemma. If $\lambda_n = 0$ for all n, then $\rho(x, S) = 0$, so $x \in S$ and we can take $s = x$; whereas if $\lambda_N = 1$, then $\rho(x, S) > 0$ and we can take $s = s_0$.

In order to obtain a constructive proof of Bishop's lemma, we replace LPO by the completeness of S. This is a typical situation, in which a result proved classically using an omniscience principle is proved constructively under some completeness condition using the λ-technique.

Now, it follows immediately from Bishop's lemma that if S is a complete, located subset of a metric space, then $\sim S = -S$. The λ-technique provides us with a related result in a Banach space, in which locatedness is replaced by convexity.

A subset C of a linear space X is said to be

▷ *convex* if $tx + (1 - t)\, y \in C$ whenever $x, y \in C$ and $0 \leqslant t \leqslant 1$;

▷ *absorbing* if for each $x \in X$ there exists $t > 0$ such that $x \in tC$.

Proposition 3.1.2. *If C is a convex, absorbing subset of a Banach space X, then $\sim C$ is dense in $\neg C$.*

Proof. Let $x \in \neg C$, let $\varepsilon > 0$, and choose $\delta > 0$ such that $\delta \|x\| < \varepsilon$. Then

$$x' = (1+\delta)x \notin (1+\delta)C.$$

Given $y \in C$, we show that $x' \neq y$. To that end, construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow \|x' - y\| < \frac{1}{(n+1)^2},$$
$$\lambda_n = 1 \Longrightarrow x' \neq y.$$

We may assume that $\lambda_1 = 0$. Define a sequence $(z_n)_{n \geqslant 1}$ in X as follows. If $\lambda_n = 0$, set $z_n = 0$; if $\lambda_n = 1 - \lambda_{n-1}$, set $z_k = n(x' - y)$ for all $k \geqslant n$. Since in the latter case we have

$$\|z_n - z_{n-1}\| = n \|x' - y\| < \frac{1}{n},$$

it follows that $\|z_m - z_n\| < 1/n$ whenever $m > n$. Hence $(z_n)_{n \geqslant 1}$ is a Cauchy sequence and so converges to a limit z in the complete space X. Choose a positive integer N such that $z \in N\delta C$, and consider any integer $n \geqslant N$. If $\lambda_n = 1 - \lambda_{n-1}$, then $z = n(x' - y)$; whence

$$x' = y + \frac{1}{n} z \in y + \frac{N\delta}{n} C \subset C + \delta C = (1+\delta)C,$$

a contradiction. Thus $\lambda_n = \lambda_{n-1}$ for all $n \geqslant N$. It follows that if $\lambda_N = 0$, then $\lambda_n = 0$ for all n, and therefore $x' = y \in C \subset (1+\delta)C$. This contradiction ensures that $\lambda_N = 1$; whence $x' \neq y$. Since $y \in C$ is arbitrary, it follows that $x' \in \sim C$. Since also $\|x' - x\| = \delta \|x\| < \varepsilon$, and $x \in \neg C$ and $\varepsilon > 0$ are arbitrary, we conclude that $\sim C$ is dense in $\neg C$. □

Classically, as the reader is invited to prove, the completeness of the space X can be removed from Proposition 3.1.2 since we are allowed to use LPO.

For an example of the use of the λ-technique to avoid the application of Markov's principle, we turn to an elementary result in metric topology. We observed in Chapter 2 that it cannot be proved constructively that the union of two closed subsets of a metric space is closed. Nevertheless, it seems reasonable to expect that if A, B are closed subsets of $\mathbb{R}$, then the complement, in some appropriate sense, of $\overline{A \cup B}$ will be the intersection of the complements of A and B. Since it is trivial that the first of these complements is a subset of the intersection of the other two, it is enough to prove the reverse inclusion. With the help of Markov's principle, we can prove that

$$(\neg A \cap \neg B) \cap \overline{A \cup B} = \varnothing.$$

Indeed, supposing that $x \in (\neg A \cap \neg B) \cap \overline{A \cup B}$, choose a sequence $(x_n)_{n \geqslant 1}$ of elements of $A \cup B$ converging to x. Define an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\begin{aligned} \lambda_n = 0 &\Longrightarrow \forall k \leqslant n \ \ (x_k \in A), \\ \lambda_n = 1 - \lambda_{n-1} &\Longrightarrow x_n \in B. \end{aligned}$$

We may assume that $\lambda_1 = 0$. If $\lambda_n = 0$ for all n, then $x \in \overline{A} = A$, a contradiction. Hence, by Markov's principle, there exists $n > 1$ such that $\lambda_n = 1 - \lambda_{n-1}$, so there exists $n_1 > 1$ with $x_{n_1} \in B$. Repeated application of this argument enables us to construct a strictly increasing sequence $(n_k)_{k \geqslant 1}$ of positive integers such that $x_{n_k} \in A$ for all even k, and $x_{n_k} \in B$ for all odd k. Thus $x \in A \cap B$, which is a final contradiction.

Since such an argument, depending on Markov's principle, is not good enough for our constructive purposes, it is fortunate that if we work within a complete metric space and use the λ-technique (on more than one occasion), we can say something interesting about the intersection of the complements of two closed sets. This will require a preliminary result.

Lemma 3.1.3. *Let X be a complete metric space, A a closed subset of X, and B a subset of X. Let $x \in {\sim}A$, let $y \in \overline{A \cup B}$, and let $(y_n)_{n \geqslant 1}$ be a sequence in $A \cup B$ that converges to y. Then either $x \neq y$ or there exists n such that $y_n \in B$.*

Proof. We may assume that $y_1 \in A$ and, by passing to a subsequence if necessary, that $\rho(y_n, y) < 1/n$ for each n. Construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\begin{aligned} \lambda_n = 0 &\Longrightarrow \forall k \leqslant n \ \ (y_k \in A), \\ \lambda_n = 1 - \lambda_{n-1} &\Longrightarrow y_n \in B. \end{aligned}$$

If $\lambda_n = 0$, set $\xi_n = y_n$; if $\lambda_n = 1 - \lambda_{n-1}$, set $\xi_k = \xi_{n-1}$ for all $k \geqslant n$. Then $(\xi_n)_{n \geqslant 1}$ is a Cauchy sequence in A; in fact,

$$\rho(\xi_m, \xi_n) \leqslant \rho(\xi_m, y) + \rho(\xi_n, y) < \frac{2}{n} \quad (m \geqslant n).$$

Since X is complete and A is closed, $(\xi_n)_{n \geqslant 1}$ converges to a limit $\xi \in A$. Choose a positive integer N such that $\rho(x, \xi) > 2/N$. If $\lambda_N = 1$, then $y_n \in B$ for some $n \leqslant N$. If $\lambda_N = 0$, then either $x \neq y$ or, as we may suppose, $\rho(x, y) < 1/N$. If there exists $m \geqslant N$ such that $\lambda_{m+1} = 1 - \lambda_m$, then

$$\begin{aligned} \rho(x, y) &\geqslant \rho(x, \xi) - \rho(\xi, y) \\ &> \frac{2}{N} - \rho(y_m, y) > \frac{2}{N} - \frac{1}{m} \geqslant \frac{1}{N}, \end{aligned}$$

a contradiction. Hence $\lambda_n = 0$ for all $n \geqslant N$ and therefore for all n; so $y_n \in A$ for all n. Since A is closed, $y \in A$ and therefore $x \neq y$. $\square$

Proposition 3.1.4. *If A, B are closed subsets of a complete metric space X, then*

$$\sim A \cap \sim B = \sim \overline{A \cup B}.$$

Proof. It is clear that $\sim\overline{A \cup B} \subset \sim A \cap \sim B$. To prove the reverse inclusion, given $x \in \sim A \cap \sim B$ and $y \in \overline{A \cup B}$, we must prove that $x \neq y$. To this end, choose a sequence $(y_n)_{n \geqslant 1}$ in $A \cup B$ that converges to y. We may assume that $y_1 \in A$ and that $\rho(y_n, y) < 1/n$ for each n. Set $\lambda_1 = 0$, $n_1 = 1$, and $\xi_1 = y_1$. In view of Lemma 3.1.3, we may also assume that there exists $n_2 > 1$ such that $y_{n_2} \in B$; set $\lambda_2 = 0$ and $\eta_1 = y_{n_2}$. Now apply the same lemma, but with the roles of A and B interchanged. Either we have $x \neq y$, when for each $k \geqslant 3$ we set $\lambda_k = 1$, $n_k = n_2$, $\xi_{k-1} = \xi_1$, and $\eta_{k-1} = \eta_1$; or else, as we may assume, there exists $n_3 > n_2$ such that $y_{n_3} \in A$, in which case we set $\lambda_3 = 0$ and $\xi_2 = y_{n_3}$. Repeating such applications of Lemma 3.1.3, we construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$, an increasing sequence $(n_k)_{k \geqslant 1}$ of positive integers, a sequence $(\xi_k)_{k \geqslant 1}$ of elements of A, and a sequence $(\eta_k)_{k \geqslant 1}$ of elements of B, such that for each $k \geqslant 1$,

- if $\lambda_{2k} = 0$, then $n_{2k} > n_{2k-1}$ and $\eta_k = y_{n_{2k}} \in B$;
- if $\lambda_{2k+1} = 0$, then $n_{2k+1} > n_{2k}$ and $\xi_{k+1} = y_{n_{2k+1}} \in A$;
- if $\lambda_{2k} = 1 - \lambda_{2k-1}$, then $x \neq y$ and for each $j \geqslant k$, $n_{2j+1} = n_{2j} = n_{2j-1}$, $\xi_j = \xi_k$, and $\eta_j = \eta_{k-1}$;
- if $\lambda_{2k+1} = 1 - \lambda_{2k}$, then $x \neq y$ and for each $j \geqslant k$, $n_{2j+2} = n_{2j+1} = n_{2k}$, $\xi_j = \xi_k$, and $\eta_j = \eta_k$.

Then $(\xi_k)_{k \geqslant 1}$ is a Cauchy sequence in A, and $(\eta_k)_{k \geqslant 1}$ is a Cauchy sequence in B. Indeed, for $m \geqslant n$ we have

$$\rho(\xi_m, \xi_n) \leqslant \rho(\xi_m, y) + \rho(\xi_n, y) < \frac{2}{n},$$

and similarly $\rho(\eta_m, \eta_n) < 2/n$. Since X is complete and the subsets A, B are closed in X, the sequences $(\xi_n)_{n \geqslant 1}, (\eta_n)_{n \geqslant 1}$ converge to respective limits $\xi \in A$ and $\eta \in B$. Choose an integer $N > 4$ such that $\rho(x, \xi) > 3/N$ and $\rho(x, \eta) > 3/N$. If $\lambda_N = 1$, then $x \neq y$; so we may assume that $\lambda_N = 0$. Either $x \neq y$ or, as we may further assume, $\rho(x, y) < 1/N$. Suppose there exists $m > N$ such that $\lambda_m = 1 - \lambda_{m-1}$. In the case where m is even, we have

$$\begin{aligned}
\rho(x, y) &\geqslant \rho(x, \eta) - \rho(\eta, y) \\
&> \frac{3}{N} - \rho\left(y_{n_{m-2}}, y\right) \\
&> \frac{3}{N} - \frac{1}{m-2} \\
&> \frac{3}{N} - \frac{2}{m} \qquad \text{(since } m > 4\text{)} \\
&> \frac{1}{N},
\end{aligned}$$

a contradiction. If m is odd, a similar argument again leads us to a contradiction. Hence $\lambda_n = 0$ for all $n \geqslant N$ and therefore for all n. It follows that $y = \xi = \eta$ and therefore $\rho(x, y) > 3/N$, a contradiction. Hence the case $\rho(x, y) < 1/N$ is ruled out, and so $\rho(x, y) \geqslant 1/N$. □

The λ-technique turns out to be very useful for clarifying the connections between various continuity properties of functions between metric spaces.

It is trivial that a continuous mapping between metric spaces is strongly extensional. If the domain is complete, then we can weaken continuity to sequential continuity.

Proposition 3.1.5. *Let f be a sequentially continuous mapping of a complete metric space X into a metric space Y. Then f is strongly extensional.*

Proof. Given points x, y of X with $f(x) \neq f(y)$, construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow \rho(x, y) < \frac{1}{n+1},$$
$$\lambda_n = 1 \Longrightarrow x \neq y.$$

Define a sequence $(z_n)_{n \geqslant 1}$ in X as follows: if $\lambda_n = 0$, set $z_n = x$; if $\lambda_n = 1 - \lambda_{n-1}$, set $z_k = y$ for all $k \geqslant n$. Then $(z_n)_{n \geqslant 1}$ is a Cauchy sequence and so converges to a limit z in X. Since f is sequentially continuous, the sequence $(f(z_n))_{n \geqslant 1}$ converges to the limit $f(z)$ in Y. Choose a positive integer N such that $\rho(f(z_n), f(z)) < \rho(f(x), f(y))$ for all $n \geqslant N$. If $\lambda_{n+1} = 1 - \lambda_n$ for some $n \geqslant N$, then $z = y$ and $z_n = x$, so

$$\rho(f(x), f(y)) = \rho(f(z_n), f(z)) < \rho(f(x), f(y)),$$

a contradiction. Hence $\lambda_{n+1} = \lambda_n$ for all $n \geqslant N$. It follows that if $\lambda_N = 0$, then $\lambda_n = 0$ for all n; whence $x = y$, and so $f(x) = f(y)$. This contradiction ensures that $\lambda_N = 1$. □

For a linear mapping on a Banach space we obtain strong extensionality without even sequential continuity, as a corollary of the following application of the λ-technique.

Proposition 3.1.6. *Let T be a linear mapping of a Banach space X into a normed space Y. Let B be a subset of the graph of T that is closed and located in $X \times Y$, and let $(x, y) \in X \times Y$ be such that $y \neq Tx$. Then $\rho((x, y), B) > 0$.*

Proof. Construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow \rho\left((x,y), B\right) < \frac{1}{(n+1)^2},$$
$$\lambda_n = 1 \Longrightarrow \rho\left((x,y), B\right) > 0.$$

We may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, set $z_n = 0$; if $\lambda_n = 1 - \lambda_{n-1}$, pick $\xi \in X$ such that

$$(\xi, T\xi) \in B \text{ and } \|x - \xi\| + \|y - T\xi\| < \frac{1}{n^2} \tag{3.1}$$

and set $z_i = n(x - \xi)$ for all $i \geqslant n$. If $j > k$ and $\|z_j - z_k\| > 0$, then $\lambda_k = 0$ and $\lambda_n = 1 - \lambda_{n-1}$ for a unique value of n with $k < n \leqslant j$; so there exists $\xi \in X$ such that (3.1) holds and

$$\|z_j - z_k\| = \|n(x - \xi)\| < \frac{1}{n} < \frac{1}{k}.$$

Hence $\|z_j - z_k\| < 1/k$ whenever $j \geqslant k$. Therefore $(z_n)_{n \geqslant 1}$ is a Cauchy sequence in X and so converges to a limit $z \in X$. Now choose a positive integer N such that

$$1 + \|Tz\| \leqslant N\|y - Tx\|$$

and consider any integer $n > N$. If $\lambda_n = 1 - \lambda_{n-1}$, then $z = n(x - \xi)$ for some ξ satisfying (3.1), so

$$\begin{aligned} 1 + \|Tz\| &= 1 + n\|Tx - T\xi\| \\ &\geqslant 1 + n\|Tx - y\| - n\|y - T\xi\| \\ &> 1 + N\|Tx - y\| - \frac{1}{n} \\ &> N\|Tx - y\|, \end{aligned}$$

a contradiction of our choice of N. Hence $\lambda_n = \lambda_{n-1}$ for all $n \geqslant N$. It follows that if $\lambda_N = 0$, then $\lambda_n = 0$ for all n, and therefore $\rho\left((x,y), B\right) = 0$. Since B is closed in $X \times Y$, we now have $y = Tx$, which contradicts our original hypotheses. We conclude that $\lambda_N = 1$. □

Corollary 3.1.7. *A linear mapping of a Banach space into a normed space is strongly extensional.*

Proof. Let T be a linear mapping of a Banach space X into a normed space Y, and let $x \in X$ satisfy $Tx \neq 0$. Given $z \in \ker(T)$, apply Proposition 3.1.6 with $y = 0$ and $B = \{(z, Tz)\}$, to obtain $(x, 0) \neq (z, Tz)$; whence either $x \neq z$ or else $Tz \neq 0$. Since the latter is absurd, we conclude that $x \neq z$. □

Markov's principle is all that we need add to intuitionistic logic in order to prove Proposition 3.1.6 without the hypothesis that X is complete: for the argument towards the end of the proof of that proposition shows that it is impossible that $\rho\left((x,y), B\right) = 0$, so, by Markov's principle, $\rho\left((x,y), B\right) > 0$. In fact (this is left to the exercises at the end of the chapter), if Proposition 3.1.6 holds without the completeness of X, then we can derive Markov's principle.

3.2 Ishihara's Tricks

In [57], Ishihara introduced the following two lemmas, now called *Ishihara's tricks*, in which we use completeness to make a decision which at first sight would seem to be impossible with purely constructive techniques.

Lemma 3.2.1. *Let f be a strongly extensional mapping of a complete metric space X into a metric space Y, and let $(x_n)_{n\geqslant 1}$ be a sequence in X converging to a limit x. Then for all positive numbers α, β with $\alpha < \beta$, either $\rho(f(x_n), f(x)) > \alpha$ for some n or $\rho(f(x_n), f(x)) < \beta$ for all n.*

Proof. Let $(n_k)_{k\geqslant 1}$ be a strictly increasing sequence of integers such that $\rho(x, x_n) < 1/(k+1)$ for all $n \geqslant n_k$. For convenience, set $n_0 = 1$. For each $k \geqslant 1$ set

$$s_k = \max\{\rho(f(x_n), f(x)) : n_{k-1} \leqslant n < n_k\}.$$

Construct an increasing binary sequence $(\lambda_k)_{k\geqslant 1}$ such that

$$\lambda_k = 0 \Longrightarrow \forall j \leqslant k\, (s_j < \beta)\,,$$
$$\lambda_k = 1 \Longrightarrow \exists j \leqslant k\, (s_j > \alpha)\,.$$

We may assume that $\lambda_1 = 0$. Define a sequence $(y_k)_{k\geqslant 1}$ in X as follows. If $\lambda_k = 0$, set $y_k = x$; if $\lambda_k = 1 - \lambda_{k-1}$, choose ν_k with $n_{k-1} \leqslant \nu_k < n_k$ and $\rho(f(x_{\nu_k}), f(x)) > \alpha$, and set $y_j = x_{\nu_k}$ for all $j \geqslant k$. Let $i \geqslant j$. If $\lambda_i = 0$ or $\lambda_j = 1$, then $y_i = y_j$. If $\lambda_i = 1 - \lambda_j$, then there exists k with $j < k \leqslant i$ and $\lambda_k = 1 - \lambda_{k-1}$; so $y_j = x$, and $y_i = y_k = x_{\nu_k}$ for some ν_k such that $n_{k-1} \leqslant \nu_k < n_k$ and $\rho(f(x_{\nu_k}), f(x)) > \alpha$; whence

$$\rho(y_i, y_j) = \rho\left(x_{\nu_k}, x\right) \leqslant \frac{1}{\nu_k + 1} < \frac{1}{j}.$$

It follows that $\rho(y_i, y_j) < 1/j$ whenever $i \geqslant j$; so $(y_k)_{k\geqslant 1}$ is a Cauchy sequence and therefore converges to a limit y in X. Either $\rho(f(x), f(y)) < \alpha$ or $\rho(f(x), f(y)) > 0$. In the first case, if $\lambda_k = 1 - \lambda_{k-1}$, then $y = x_{\nu_k}$ with $\rho(f(x_{\nu_k}), f(x)) > \alpha$, a contradiction; whence $\lambda_k = 0$ for all k. Then $\rho(f(x_n), f(x)) < \beta$ for all n. In the second case, since f is strongly extensional, $x \neq y$. Choose a positive integer κ such that $x \neq y_\kappa$. If $\lambda_\kappa = 0$, then $x \neq y_\kappa = x$, a contradiction; whence $\lambda_\kappa = 1$ and there exists n such that $\rho(f(x_n), f(x)) > \alpha$. □

Let $P(n)$ be a property of positive integers n. We say that $P(n)$ holds

- *eventually*, or *for all sufficiently large n*, if there exists N such that $P(n)$ holds for all $n \geqslant N$;
- *infinitely often* if for each n there exists $m > n$ such that $P(m)$ holds.

The foregoing lemma leads to a technique allowing us to decide between alternatives that happen eventually and those that happen infinitely often.

Lemma 3.2.2. ***(Ishihara's second trick)*** *Let f be a strongly extensional mapping of a complete metric space X into a metric space Y, and let $(x_n)_{n\geqslant 1}$ be a sequence in X converging to a limit x. Then for all positive numbers α, β with $\alpha < \beta$, either $\rho(f(x_n), f(x)) > \alpha$ for infinitely many n or else $\rho(f(x_n), f(x)) < \beta$ for all sufficiently large n.*

Proof. Let $(n_k)_{k\geqslant 1}$ be a strictly increasing sequence of positive integers such that $\rho(x, x_n) < 1/k$ for all $n \geqslant n_k$. Successively applying Lemma 3.2.1 to the subsequence $(x_n)_{n\geqslant n_k}$, construct an increasing binary sequence $(\lambda_k)_{k\geqslant 1}$ such that

$$\lambda_k = 0 \Longrightarrow \exists n \geqslant n_k \left(\rho(f(x_n), f(x)) > \alpha\right),$$
$$\lambda_k = 1 \Longrightarrow \forall n \geqslant n_k \left(\rho(f(x_n), f(x)) < \beta\right).$$

We may assume that $\lambda_1 = 0$. Define a sequence $(y_k)_{k\geqslant 1}$ in X as follows. If $\lambda_k = 0$, choose $\nu_k \geqslant n_k$ such that $\rho(f(x_{\nu_k}), f(x)) > \alpha$, and set $y_k = x_{\nu_k}$; if $\lambda_k = 1-\lambda_{k-1}$, set $y_i = y_{k-1}$ for all $i \geqslant k$. Then $(y_k)_{k\geqslant 1}$ is a Cauchy sequence: in fact, $\rho(y_i, y_j) < 2/j$ whenever $i \geqslant j$. Let y be the limit of $(y_k)_{k\geqslant 1}$ in the complete space X. Either $0 < \rho(f(x), f(y))$ or $\rho(f(x), f(y)) < \alpha$. In the first case, since f is strongly extensional, $x \neq y$, so there exists a positive integer κ with $\rho(x, y_\kappa) > 1/\kappa$. If $\lambda_\kappa = 0$, then $y_\kappa = x_{\nu_\kappa}$ for some $\nu_\kappa \geqslant n_\kappa$; whence $\rho(x, y_\kappa) < 1/\kappa$, a contradiction. Thus $\lambda_\varkappa = 1$, and therefore $\rho\left(f(x_n), f(x)\right) < \beta$ eventually. In the case $\rho(f(x), f(y)) < \alpha$, if there exists k such that $\lambda_{k+1} = 1 - \lambda_k$, then $y = x_{\nu_k}$ and $\rho(f(x_{\nu_k}), f(x)) > \alpha$, a contradiction; whence $\lambda_k = 0$ for all k, and therefore $\rho(f(x_n), f(x)) > \alpha$ infinitely often. □

A mapping $f : X \longrightarrow Y$ between metric spaces is said to be *sequentially nondiscontinuous* if it has the following property: if $(x_n)_{n\geqslant 1}$ converges to $x \in X$ and $\rho\left(f\left(x_n\right), f(x)\right) \geqslant \delta$ for all n, then $\delta \leqslant 0$. Clearly, a sequentially continuous mapping is sequentially nondiscontinuous. Ishihara's first application of his tricks was to the converse.

Proposition 3.2.3. *Let f be a mapping of a complete metric space X into a metric space Y. Then f is sequentially continuous if and only if it is strongly extensional and sequentially nondiscontinuous.*

Proof. Suppose that f is strongly extensional and sequentially nondiscontinuous. Let $(x_n)_{n\geqslant 1}$ be a sequence converging to $x \in X$, and let $\varepsilon > 0$. By Lemma 3.2.2, either there exists a subsequence $(x_{n_k})_{k\geqslant 1}$ of $(x_n)_{n\geqslant 1}$ such that $\rho\left(f(x_{n_k}), f(x)\right) > \varepsilon/2$ for all k, or else $\rho\left(f(x_n), f(x)\right) < \varepsilon$ for all sufficiently large n. In the former case, the sequential nondiscontinuity of f shows that $\varepsilon \leqslant 0$, which is absurd. We conclude that the latter case obtains. Hence f is sequentially continuous.

The converse is an immediate consequence of Proposition 3.1.5 and the observation made before this proposition. □

Corollary 3.2.4. *For linear mappings of a Banach space into a normed space, sequential continuity and sequential nondiscontinuity are equivalent.*

Proof. This follows immediately from Corollary 3.1.7 and Proposition 3.2.3. □

We now lift Ishihara's tricks into a general setting. This both clarifies the ideas underlying those lemmas and raises the possibility that some other applications of their proof techniques in constructive analysis are, in fact, corollaries of our general results. We begin with a generalisation of Ishihara's first trick (Lemma 3.2.1).

Proposition 3.2.5. *Let X be a complete metric space, let P, Q be subsets of X such that $X = P \cup Q$, and let x be an element of X such that for each $y \in X$, either $x \neq y$ or $y \notin Q$. Then for each sequence $(x_n)_{n\geqslant 1}$ in X that converges to x, either $x_n \in P$ for all n or else there exists N such that $x_N \in Q$.*

Proof. Construct an increasing binary sequence $(\lambda_n)_{n\geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow \forall k \leqslant n \, (x_k \in P),$$
$$\lambda_n = 1 - \lambda_{n-1} \Longrightarrow x_n \in Q.$$

We may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, set $y_n = x$; if $\lambda_n = 1-\lambda_{n-1}$, set $y_k = x_n$ for all $k \geqslant n$. To see that $(y_n)_{n\geqslant 1}$ is a Cauchy sequence in X, let $\varepsilon > 0$ and compute N such that $\rho(x, x_n) < \varepsilon$ for all $n \geqslant N$. Let $m \geqslant n \geqslant N$. If $\lambda_m = 0$, then $y_m = y_n = x$. So we may assume that $\lambda_m = 1$. If $\lambda_n = 0$, then

$$\rho(y_m, y_n) = \rho(x_m, x) < \varepsilon.$$

If $\lambda_n = 1$, then

$$\rho(y_m, y_n) = \rho(x_m, x_n) \leqslant \rho(x, x_m) + \rho(x, x_n) < 2\varepsilon.$$

Thus in all cases, $\rho(y_m, y_n) < 2\varepsilon$.

Since X is complete, the Cauchy sequence $(y_n)_{n\geqslant 1}$ converges to a limit $y \in X$. Either $x \neq y$ or $y \notin Q$. In the first case, choosing N such that $x \neq y_N$, we see that $\lambda_N = 1$. In the second case, if there exists m such that $\lambda_m = 1 - \lambda_{m-1}$, then $y = x_m \in Q$, a contradiction; whence $\lambda_n = 0$, and therefore $x_n \in P$, for all n. □

To derive Lemma 3.2.1, assume the hypotheses of that lemma, define

$$P = \{y \in X : \rho(f(y), f(x)) < \beta\},$$
$$Q = \{y \in X : \rho(f(y), f(x)) > \alpha\},$$

and apply Proposition 3.2.5, noting that for all $y \in Y$,

- either $\rho(f(y), f(x)) > 0$, in which case, by the strong extensionality of f, $x \neq y$;
- or else $\rho(f(y), f(x)) < \alpha$ and therefore $y \notin Q$.

In order to tackle the generalisation of Ishihara's second trick (Lemma 3.2.2), we prove two lemmas, the first of which is a variant of Proposition 3.2.5.

Lemma 3.2.6. *Let X be a complete metric space, let P, Q be subsets of X such that $X = P \cup Q$, and let x be a point of X such that for each $y \in X$, either $x \neq y$ or $y \notin P$. Then for each sequence $(x_n)_{n \geqslant 1}$ in X that converges to x, either $x_n \in P$ for all n or else there exists N such that $x_N \in Q$.*

Proof. Construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow \forall k \leqslant n \ (x_k \in P),$$
$$\lambda_n = 1 - \lambda_{n-1} \Longrightarrow x_n \in Q.$$

We may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, set $y_n = x$; if $\lambda_n = 1 - \lambda_{n-1}$, set $y_k = x_{n-1}$ for each $k \geqslant n$. Then $(y_n)_{n \geqslant 1}$ is a Cauchy sequence in X and so converges to a limit $y \in X$. Either $x \neq y$ or else $y \notin P$. In the first case choose N such that $\rho(x_{n-1}, y) > \frac{1}{2}\rho(x, y)$ for all $n \geqslant N$, and suppose that $\lambda_N = 0$. If $\lambda_m = 1 - \lambda_{m-1}$ for some $m > N$, then $y = x_{m-1}$ and so $0 = \rho(x_{m-1}, y) > \frac{1}{2}\rho(x, y)$, a contradiction. Hence $\lambda_n = 0$ for all $n \geqslant N$ and therefore for all n; but this implies that $y = x$, another contradiction. Thus, in fact, $\lambda_N = 1$ and there exists $n \leqslant N$ such that $x_n \in Q$. On the other hand, in the case $y \notin P$ we must have $\lambda_n = 0$, and therefore $x_n \in P$, for all n. □

Lemma 3.2.7. *Let X be a complete metric space, let P, Q be subsets of X such that $X = P \cup Q$, and let x be an element of X. Suppose that for any sequence $(x_n)_{n \geqslant 1}$ converging to x in X, either there exists N such that $x \neq x_N$ or else $x_n \notin Q$ for all n. Then for each $y \in X$, either $x \neq y$ or $y \notin Q$.*

Proof. Applying the hypotheses to the sequence $(x, x, \ldots)$ in X, we see that $x \notin Q$. Given y in X, construct a (perforce increasing) binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow \rho(x, y) < \frac{1}{n},$$
$$\lambda_n = 1 \Longrightarrow \rho(x, y) > \frac{1}{n+1}.$$

We may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, set $y_n = x$; if $\lambda_n = 1 - \lambda_{n-1}$, set $y_n = y$ and $y_k = x$ for each $k \geqslant n+1$. It is easily shown that the sequence $(y_n)_{n \geqslant 1}$ converges to x; so either there exists N such that $x \neq y_N = y$, or else $y_n \notin Q$ for each n. In the latter case, supposing that $y \in Q$, we see that $\lambda_n = 0$ for all n; whence $y = x$, which is impossible, since we have already shown that $x \notin Q$. We conclude that $y \notin Q$. □

This brings us to the generalisation of Ishihara's second trick (Lemma 3.2.2).

Proposition 3.2.8. *Let X be a complete metric space, let P, Q be subsets of X such that $X = P \cup Q$, and let x be an element of X. Suppose that for any sequence $(x_n)_{n\geqslant 1}$ converging to x in X, either there exists N such that $x \neq x_N$ or else $x_n \notin Q$ for all n. Then for any such sequence, either $x_n \in P$ eventually or else $x_n \in Q$ infinitely often.*

Proof. Consider the space $\mathbf{X}$ of all sequences that converge to x in X, taken with the metric

$$\rho(\boldsymbol{\xi}, \boldsymbol{\eta}) = \sup_{n\geqslant 1} \rho(\xi_n, \eta_n),$$

where, for example, $\boldsymbol{\xi}$ denotes the sequence $(\xi_1, \xi_2, \ldots)$ in $\mathbf{X}$. An argument like the standard one for proving the completeness of the classical sequence space l_∞ (and left as an exercise) shows that $\mathbf{X}$ is complete. Define

$$\mathbf{P} = \{\mathbf{x} \in \mathbf{X} : \exists n\,(x_n \in Q)\},$$
$$\mathbf{Q} = \{\mathbf{x} \in \mathbf{X} : \forall n\,(x_n \in P)\}.$$

We see from our hypotheses and Lemma 3.2.7 that for each $y \in X$, either $x \neq y$ or $y \notin Q$. It follows from Proposition 3.2.5 that for any sequence $\mathbf{x}$ in X that converges to x, either there exists n such that $x_n \in Q$ or else $x_n \in P$ for all n; whence either $\mathbf{x} \in \mathbf{P}$ or $\mathbf{x} \in \mathbf{Q}$. Let $\mathbf{x}^\infty = (x, x, x, \ldots)$, and for each n let $\mathbf{x}^n = (x_n, x_{n+1}, \ldots)$. Then $\mathbf{x}^n \longrightarrow \mathbf{x}^\infty$ in $\mathbf{X}$. Since our hypotheses also ensure that for each $\mathbf{y} \in \mathbf{X}$, either $\mathbf{x}^\infty \neq \mathbf{y}$ or $\mathbf{y} \notin \mathbf{P}$, we can apply Lemma 3.2.6 in the space $\mathbf{X}$ to show that either $\mathbf{x}^n \in \mathbf{P}$ for each n or else there exists N such that $\mathbf{x}^N \in \mathbf{Q}$. This, when unwrapped, is precisely the conclusion we want. □

It is relatively simple to prove Lemma 3.2.2 using Proposition 3.2.8. Under the hypotheses of that lemma, take

$$P = \{y \in X : \rho(f(x), f(y)) < \beta\},$$
$$Q = \{y \in X : \rho(f(x), f(y)) > \alpha\},$$

and, for convenience,

$$P' = \{y \in X : \rho(f(x), f(y)) < \alpha\},$$
$$Q' = \{y \in X : \rho(f(x), f(y)) > \alpha/2\}.$$

Then $X = P \cup Q$ and for each $y \in X$, either $x \neq y$ or $y \notin Q$. Similarly, $X = P' \cup Q'$ and for each $y \in X$, either $x \neq y$ or $y \notin Q'$. It follows from the last statement and Proposition 3.2.5 that if $(y_n)_{n\geqslant 1}$ is any sequence converging to x in X, then either $y_n \in P'$ for all n or else there exists N such that $y_N \in Q'$. Hence either $y_n \notin Q$ for all n or else there exists N such that $x \neq y_N$. Thus P and Q satisfy the hypotheses of Proposition 3.2.8, from which we immediately deduce the conclusion of Ishihara's second trick.

We have now given a number of illustrations of the merit of the λ-technique as a replacement for omniscience principles. The technique will reappear throughout the book, where it will be used, for example, to construct best approximations from finite-dimensional spaces, to locate certain convex sets in a normed space, and to discuss the range of an operator with an adjoint on a Hilbert space.

Exercises

1. Using the λ-technique, show that if $\mathbb{R}a = \{ax : x \in \mathbb{R}\}$ is closed, then $a = 0$ or $a \neq 0$.

2. Let X be a complete metric space, and $f : X \longrightarrow \mathbb{R}$ a sequentially continuous mapping such that $\inf f$ exists. Suppose that to each $\varepsilon > 0$ there corresponds $\delta > 0$ such that if $x, y \in X$ and $\max\{f(x), f(y)\} < \delta$, then $\rho(x, y) < \varepsilon$. Prove that there exists $a \in X$ such that if $f(a) > 0$, then $\inf f > 0$.

3. A linear mapping $T : X \longrightarrow Y$ between normed spaces is said to be *well-behaved* if $Tx \neq 0$ whenever $x \in {\sim}\ker(T)$. Prove that the statement "every linear mapping between normed spaces is well-behaved" is equivalent to Markov's principle. Prove also that every linear mapping from a Banach space into a normed space is well-behaved.

4. Prove that if Corollary 3.1.7 (and hence a fortiori Proposition 3.1.6) holds without the hypothesis that X is complete, then we can derive Markov's principle.

5. Prove *de Morgan's rule for metric complements*: If $(S_n)_{n\geqslant 1}$ is a sequence of located subsets of a metric space X such that $S = \bigcup_{n\geqslant 1} S_n$ is complete and located, then $-S = \bigcap_{n\geqslant 1} -S_n$.

6. A subset S of a metric space X is said to be *uniformly almost located* if there exists a strictly decreasing sequence $(\delta_n)_{n\geqslant 1}$ of positive numbers converging to 0 such that the following holds: for each $x \in X$ there exists $y \in S$ such that for each n, if $\rho(x, y) > \delta_n$, then $\rho(x, S) > \delta_{n+1}$. Prove that if S is an inhabited, uniformly almost located subset of a locally totally bounded space, then S is located.

7. Let X be a complete metric space, and a a point of X such that
$$s = \sup\{\rho(a, x) : x \in X\}$$
exists. Prove that for each $r > 1$ there exists $b \in X$ such that $s \leqslant r\rho(a, b)$.

8. Let T be a sequentially continuous linear mapping of a Banach space X into a normed space Y, and let B be the unit ball of the range of T. Suppose that for each $\varepsilon > 0$, either there exists $x \in {\sim}T^{-1}(B)$ with $\|x\| < \varepsilon$ or else ${\sim}T^{-1}(B)$ is bounded away from 0 (that is, there exists $\delta > 0$ such that $\|x\| > \delta$ whenever $T(x) \in B$). Prove that T is a bounded linear mapping.

9. Let $T : X \longrightarrow Y$ be a sequentially continuous linear map between normed spaces such that $\ker(T)$ is located. Prove that if $x_0 \in X$ and $T(x_0) \neq 0$, then $\rho(x_0, \ker(T)) > 0$.

10. Let X be a metric space, and let S, T be subsets of X such that for each $s \in S$, each $t \in T$, and each $\varepsilon > 0$, either $s \neq t$ or there exists $y \in S \cap T$ with $\rho(t, y) < \varepsilon$. Prove that if $S \cap T$ is complete, then S and T *intersect sharply* in the following sense: if $x \in {\sim}(S \cap T)$, then for each $s \in S$ and each $t \in T$, either $x \neq s$ or $x \neq t$.

11. A sequence $(x_n)_{n \geqslant 1}$ in a metric space is said to be *weakly discriminating* if for all positive a, b with $a < b$, either $\rho(x_n, x_1) < b$ for all n or else $\rho(x_n, x_1) > a$ for some n. Prove that every totally bounded sequence in a metric space is weakly discriminating.

    Let $f : X \longrightarrow Y$ be a function between metric spaces that maps Cauchy sequences to weakly discriminating sequences. Prove that f is strongly extensional.

12. A sequence $(x_n)_{n \geqslant 1}$ in a metric space is called an *LEM-Cauchy sequence* if

$$\neg\neg\left((x_n)_{n\geqslant 1} \text{ is a Cauchy sequence}\right).$$

    Prove that the following are equivalent conditions on a mapping f between metric spaces:

    (a) f maps convergent sequences to weakly discriminating LEM-Cauchy sequences.

    (b) f is sequentially continuous.

13. Let f be a mapping of a complete metric space X into a metric space Y. Suppose that f is strongly extensional and that it is *sequentially discontinuous* at some point $x \in X$ in the following sense: there exist $\varepsilon > 0$ and a sequence $(x_n)_{n \geqslant 1}$ converging to x such that $\rho(f(x), f(x_n)) > \varepsilon$ for each n. Prove that LPO holds.

14. Let f be a strongly extensional mapping of a complete metric space X into a compact metric space Y, and let $(x_n)_{n \geqslant 1}$ be a sequence converging to x in X. Prove that the sequence $(f(x_n))_{n \geqslant 1}$ has a convergent subsequence. (*Hint:* First prove that if LPO holds, then every sequence in a compact metric space has a convergent subsequence. Then use Ishihara's tricks and Exercise 13.)

15. Let $(U_n)_{n\geqslant 1}$ be a sequence of located open sets in a complete metric space X such that the metric complement of $U = \bigcap_{n\geqslant 1} U_n$ is inhabited. Prove that there exist a point x_∞ in $-U$ and an increasing binary sequence $(\lambda_n)_{n\geqslant 1}$ such that for each n, if $\lambda_n = 0$, then $x_\infty \in U_n$, and if $\lambda_n = 1$, then $-U_k$ is inhabited for some $k \leqslant n$.

    Use this result to prove that Markov's principle is equivalent to the following statement (which is classically equivalent to Baire's theorem): If $(U_n)_{n\geqslant 1}$ is a sequence of located open subsets of a complete metric space X such that $-\bigcap_{n\geqslant 1} U_n$ is inhabited, then there exists n such that $-U_n$ is inhabited.

Notes

A variation of the argument at the beginning of the chapter shows that if for each $z \in \mathbb{C}$, there exists $\theta \in [0, 2\pi)$ such that $z = |z|\, e^{i\theta}$, then LLPO holds.

The classical proposition "sequential continuity implies pointwise continuity" is equivalent to the essentially nonconstructive principle BD-$\mathbb{N}$ that we discuss later, in Section 6.3.

When taken with the Church–Markov–Turing thesis, Exercise 13 shows that it is impossible that there exist a sequentially discontinuous function from $\mathbb{R}$ to $\mathbb{R}$, since LPO is false in RUSS. It is, however, a far cry from showing that the existence of a sequentially discontinuous mapping on $\mathbb{R}$ implies LPO to proving that every mapping $f : \mathbb{R} \longrightarrow \mathbb{R}$ is pointwise continuous. The latter can be done with the aid of either Brouwer's continuity principle or else both the Church–Markov–Turing thesis and Markov's principle; see Chapters 3 and 5 of [34].

In connection with Exercise 1, Fred Richman has shown us the following choice-free proof that if $\mathbb{R}a$ is complete, then either $a = 0$ or $a \neq 0$. Assume that $\mathbb{R}a$ is complete. For each $\varepsilon > 0$ we have either $\sqrt{|a|} < \varepsilon$ or else $|a| > 0$; in the latter case,

$$\sqrt{|a|} = \pm\frac{a}{\sqrt{|a|}} \in \mathbb{R}a.$$

Since $\varepsilon > 0$ is arbitrary, $\sqrt{|a|} \in \overline{\mathbb{R}a} = \mathbb{R}a$ and there exists r such that $\sqrt{|a|} = ra$. Pick a positive integer $N > r$. Either $|a| > 0$ or $|a| < 1/N^2$. In the latter case, if $a \neq 0$, then

$$|r|\,|a| = |ra| = \sqrt{|a|} = \frac{|a|}{\sqrt{|a|}} > N\,|a|\,,$$

so $|r| > N$, a contradiction; whence $a = 0$.

Exercise 3 and an extended version of Proposition 3.1.6 originated in [24].

Exercise 6 deals with a weak converse of Bishop's lemma. There seems to be no obvious use for such converses, in spite of the extreme value of Bishop's lemma itself.

The notion of "weakly discriminating", and Exercises 11 and 12, come from [37].

Exercise 14 produces a constructive substitute for the highly nonconstructive Bolzano–Weierstrass theorem; it first appeared in [30]. For more constructive analyses of the Bolzano–Weierstrass theorem see [27] and [22].

For more on Baire's theorem, see [34] (Chapter 2) and Chapter 6 below.

4

Finite-Dimensional and Hilbert Spaces

A mind that is stretched by a new experience can never go back to its old dimensions.

—Oliver Wendell Holmes, attrib.

We first examine finite-dimensional spaces, including an application of the λ-technique to the problem of finding best approximations by elements of a finite-dimensional subspace. We then introduce Hilbert spaces, which are natural generalisations of finite-dimensional Euclidean spaces.

4.1 Finite-Dimensional Spaces

Let X be a linear space equipped with a compatible inequality. Finitely many vectors $e_1, \ldots, e_n$ in X are said to be *linearly independent* if $\sum_{i=1}^{n} \lambda_i e_i \neq 0$ for all scalars $\lambda_1, \ldots, \lambda_n$ such that $\sum_{i=1}^{n} |\lambda_i| > 0$. In that case, if the λ_i are scalars such that $\sum_{i=1}^{n} \lambda_i e_i = 0$, then $\lambda_i = 0$ for each i. We say that the space X is *finite-dimensional* if either $X = \{0\}$ or else it contains finitely many linearly independent vectors $e_1, \ldots, e_n$ such that for each $x \in X$ there exist scalars $\lambda_1, \ldots, \lambda_n$ for which

$$x = \sum_{i=1}^{n} \lambda_i e_i. \tag{4.1}$$

In the first case we say that X has *dimension* 0 or is 0-*dimensional*. In the second case we say that X has *dimension* n or is n-*dimensional*, we call $\{e_1, \ldots, e_n\}$ a *basis* of/for X, and we say that the space X is *spanned* by, or is the (linear) *span* of, the set $\{e_1, \ldots, e_n\}$. We denote the dimension of a finite-dimensional space X

by $\dim(X)$. Since any two bases of X have the same number of vectors, $\dim(X)$ is well defined.

If $\{e_1, \ldots, e_n\}$ is a basis for X, then for each $x \in X$, the *coordinates* λ_i in the representation (4.1) are uniquely defined; so there are well-defined, clearly linear, *coordinate functionals* $u_i : X \longrightarrow \mathbb{K}$ such that $x = \sum_{i=1}^{n} u_i(x)e_i$ for each x in X. We shall prove that each of the coordinate functionals is a bounded linear map relative to any norm on X.

Lemma 4.1.1. *Let e be a nonzero vector in a normed space X. Then the subspace $Y = \mathbb{K}e$ is locally compact, and $\sim Y = -Y$.*

Proof. The mapping $\lambda \rightsquigarrow \lambda \left\|e\right\|^{-1} e$ is an isometric isomorphism of $\mathbb{K}$ onto Y, so Y is locally compact and therefore located. Since locally compact spaces are complete, it follows from Bishop's lemma (Proposition 3.1.1) that $\sim Y = -Y$. □

Lemma 4.1.2. *If Y is a finite-dimensional subspace of a normed space X, then Y is located and $\sim Y = -Y$.*

Proof. The case $\dim(Y) = 0$ is trivial, and the case $\dim(Y) = 1$ is dealt with in Lemma 4.1.1. Assume that all subspaces of dimension at most $n \geqslant 1$ in all normed spaces are located, and that their complements and metric complements coincide. Consider an $(n+1)$-dimensional space Y with basis vectors $e_1, \ldots, e_{n+1}$. Let Z be the n-dimensional space spanned by $\{e_1, \ldots, e_n\}$. First note that for all $\lambda_1, \ldots, \lambda_n$ in $\mathbb{K}$, since $1 + \sum_{i=1}^{n} |\lambda_i| > 0$, the linear independence of the vectors $e_1, \ldots, e_{n+1}$ gives $e_{n+1} \neq \sum_{i=1}^{n} \lambda_i e_i$. Hence e_{n+1} belongs to $\sim Z$ and therefore, by our induction hypothesis, to $-Z$; whence

$$\|e_{n+1}\|_{X/Z} = \rho\left(e_{n+1}, Z\right) > 0.$$

It follows that $\mathbb{K}e_{n+1}$ is a 1-dimensional subspace of X/Z. Now, for each $x \in X$,

$$\begin{aligned}\rho\left(x, Y\right) &= \inf\left\{\|x - te_{n+1} - z\| : t \in \mathbb{K},\ z \in Z\right\} \\ &= \inf\left\{\rho\left(x - te_{n+1}, Z\right) : t \in \mathbb{K}\right\} \\ &= \inf\left\{\|x - te_{n+1}\|_{X/Z} : t \in \mathbb{K}\right\},\end{aligned}$$

which exists, since, by Lemma 4.1.1, the 1-dimensional subspace $\mathbb{K}e_{n+1}$ is located in X/Z. Hence Y is located in X.

Now consider an element x of $\sim Y$. For all $t \in \mathbb{K}$ and $z \in Z$ we have $x \neq te_{n+1} + z$ and therefore $x - te_{n+1} \neq z$. Thus $x - te_{n+1} \in \sim Z$, and therefore, by our induction hypothesis,

$$\|x - te_{n+1}\|_{X/Z} = \rho\,(x - te_{n+1}, Z) > 0.$$

This shows that $x \neq_{X/Z} \lambda e_{n+1}$; whence x is in the complement of $\mathbb{K}e_{n+1}$ relative to the quotient norm on X/Z. Again applying our induction hypothesis, this time to the 1-dimensional subspace $\mathbb{K}e_{n+1}$ of X/Z, and denoting by $\rho_{X/Z}$ the distance corresponding to the quotient norm on X/Z, we see that

$$0 < \rho_{X/Z}\,(x, \mathbb{K}e_{n+1}) = \rho\,(x, Y)\,,$$

so $x \in -Y$. Hence $\sim Y \subset -Y$ and therefore $\sim Y = -Y$. This completes the induction step. □

Proposition 4.1.3. *Let $\{e_1, \ldots, e_n\}$ be a basis for an n-dimensional normed space X. Then the corresponding coordinate functionals are bounded linear functionals.*

Proof. In the case $n = 1$, the sole coordinate functional is the linear mapping $\lambda e_1 \rightsquigarrow \lambda$, which, being an isometry (that is, distance preserving), is trivially continuous. Now consider the case $n \geqslant 2$. For each $x \in X$ write $x = \sum_{i=1}^{n} u_i(x)e_i$. In order to prove that the coordinate functional u_k is bounded, we may relabel, if necessary, to take $k = n$. Let Z be the $(n-1)$-dimensional subspace of X with basis $\{e_1, \ldots, e_{n-1}\}$. Since, as in the proof of Lemma 4.1.2, $e_n \in \sim Z$, it follows from that lemma that $\rho\,(e_n, Z) > 0$ and hence that $\mathbb{K}e_n$ is a 1-dimensional subspace of X/Z. Moreover, for each $x \in X$ we have

$$\begin{aligned}
\|x\|_{X/Z} &= \rho\left(\sum_{i=1}^{n} u_i(x)e_i, Z\right) = \rho\,(u_n(x)e_n, Z) \\
&= \|u_n(x)e_n\|_{X/Z} = |u_n(x)|\,\|e_n\|_{X/Z}
\end{aligned}$$

and therefore

$$|u_n(x)| = \frac{1}{\|e_n\|_{X/Z}}\,\|x\|_{X/Z} \leqslant \frac{1}{\|e_n\|_{X/Z}}\,\|x\|\,.$$

Thus $1/\,\|e_n\|_{X/Z}$ is a bound for the linear functional u_n. □

Corollary 4.1.4. *Every linear mapping from a finite-dimensional normed space into a normed space is bounded.*

Proof. Let u be a linear mapping of a finite-dimensional normed space X into a normed space Y. We may assume that $\dim(X) > 0$. Let $\{e_1, \ldots, e_n\}$ be a basis of X, and, using Proposition 4.1.3, compute a common bound $c > 0$ for the corresponding coordinate functionals u_i $(1 \leqslant i \leqslant n)$. For each $x \in X$ we have

$$\begin{aligned}\|u(x)\| &= \left\| u\left(\sum_{i=1}^{n} u_i(x)e_i\right)\right\| = \left\|\sum_{i=1}^{n} u_i(x)u(e_i)\right\| \\ &\leqslant \sum_{i=1}^{n} \|u_i(x)u(e_i)\| = \sum_{i=1}^{n} |u_i(x)|\,\|u(e_i)\| \\ &\leqslant \left(c\sum_{i=1}^{n} \|u(e_i)\|\right)\|x\|.\end{aligned}$$

So u is a bounded linear map. □

Corollary 4.1.5. *Let $\{e_1, \ldots, e_n\}$ be a basis for an n-dimensional normed space X. Then the* **canonical mapping**

$$(\lambda_1, \ldots, \lambda_n) \rightsquigarrow \sum_{i=1}^{n} \lambda_i e_i \tag{4.2}$$

is a bounded linear injection of the Euclidean space $\mathbb{K}^n$ onto X, and its inverse is a bounded linear injection.

Proof. It is routine to verify the linearity of this map and its inverse, and that the maps are injective. The continuity of the mappings follows immediately from Corollary 4.1.4. □

Proposition 4.1.6. *A finite-dimensional normed space is locally compact.*

Proof. The 0-dimensional case is trivial. Consider a finite-dimensional normed space X with a basis $\{e_1, \ldots, e_n\}$. Let $u : \mathbb{K}^n \longrightarrow X$ be the bounded linear mapping defined at (4.2). By Corollary 4.1.5, u^{-1} is a bounded linear map. Hence for any bounded subset B of X, $u^{-1}(B)$ is a bounded subset of the locally compact space $\mathbb{K}^n$ and is therefore contained in a compact subset K of $\mathbb{K}^n$. Since u is injective and its inverse mapping is uniformly continuous, it is strongly injective. It follows from Proposition 2.2.16 that $u(K)$, which clearly contains B, is a compact subset of X. □

Corollary 4.1.7. *The unit ball of a finite-dimensional normed space is compact.*

Proof. Let X be a finite-dimensional normed space, and B its unit ball. By Proposition 4.1.6, there exists a compact subset K of X that contains B. Now, for each $x \in X$,

$$\rho(x, B) = \max\{0, 1 - \|x\|\}$$

exists, so B is totally bounded (by Proposition 2.2.10). Being also closed in K, it is complete and therefore compact. □

Corollary 4.1.8. *Every linear mapping from a finite-dimensional normed space into a normed space is normed.*

Proof. Let u be a linear mapping from a finite-dimensional normed space X into a normed space Y. We may assume that $\dim(X) \geqslant 1$. By Corollary 4.1.4, u is a bounded linear mapping and hence is uniformly continuous on X. In particular, this implies that u maps the compact (by Corollary 4.1.7) unit ball of X onto a totally bounded subset of Y; so

$$\|u\| = \sup\{\|u(x)\| : x \in X,\ \|x\| \leqslant 1\}$$

exists. □

Two norms $\|\ \|, \|\ \|'$ on a vector space X are said to be *equivalent* if both the identity mapping from $(X, \|\ \|)$ onto $(X, \|\ \|')$ and its inverse are continuous; since those mappings are linear, it follows from Proposition 2.3.3 that $\|\ \|$ and $\|\ \|'$ are equivalent norms on X if and only if there exist positive constants a, b such that $a\|x\| \leqslant \|x\|' \leqslant b\|x\|$ for all $x \in X$.

Corollary 4.1.9. *Any two norms on a finite-dimensional space are equivalent.*

Proof. If $\|\ \|$ and $\|\ \|'$ are two norms on a finite-dimensional space X, then, by Corollary 4.1.4, both the identity mapping from $(X, \|\ \|)$ to $\left(X, \|\ \|'\right)$ and its inverse are bounded, and hence continuous, linear mappings. □

We want to prove the converse of Corollary 4.1.7. This requires three lemmas, the second of which will have several applications later in the book.

Lemma 4.1.10. *Let X be a normed space, Y an n-dimensional subspace of X with basis $\{e_1, \dots, e_n\}$, and e a vector such that $\rho(e, Y) > 0$. Then the linear span of $Y \cup \{e\}$ is $(n+1)$-dimensional, with basis $\{e_1, \dots, e_n, e\}$.*

Proof. We need only prove that the vectors $e_1, \dots, e_n, e$ are linearly independent. To this end, consider elements λ_i $(1 \leqslant i \leqslant n+1)$ of $\mathbb{K}$ such that $\sum_{i=1}^{n+1} |\lambda_i| > 0$. Either $\lambda_{n+1} \neq 0$ or $\sum_{i=1}^{n} |\lambda_i| > 0$. In the first case,

$$\left\| \sum_{i=1}^{n} \lambda_i e_i + \lambda_{n+1} e \right\| \geqslant |\lambda_{n+1}|\, \rho(e, Y) > 0. \tag{4.3}$$

In the second case, since $e_1, \dots, e_n$ are linearly independent, $\left\| \sum_{i=1}^{n} \lambda_i e_i \right\| > 0$. Hence either $\|\lambda_{n+1} e\| > 0$, so that $\lambda_{n+1} \neq 0$ and we have (4.3); or else $\|\lambda_{n+1} e\| < \left\| \sum_{i=1}^{n} \lambda_i e_i \right\|$ and therefore

$$\left\| \sum_{i=1}^{n} \lambda_i e_i + \lambda_{n+1} e \right\| \geqslant \left\| \sum_{i=1}^{n} \lambda_i e_i \right\| - \|\lambda_{n+1} e\| > 0.$$

Thus in all cases we have $\sum_{i=1}^{n} \lambda_i e_i + \lambda_{n+1} e \neq 0$. □

Lemma 4.1.11. *Let S be the span of a finitely enumerable set $\{x_1, \ldots, x_n\}$ in a normed space X, and let $\varepsilon > 0$. Then there exists a finite-dimensional subspace Y of S such that $\rho(x_i, Y) < \varepsilon$ for each i.*

Proof. Setting $X_0 = \{0\}$, suppose that for some $k < n$ we have constructed finite-dimensional subspaces $X_0 \subset X_1 \subset \cdots \subset X_k \subset S$ such that $\rho(x_i, X_k) < \varepsilon$ for $1 \leqslant i \leqslant k$. Then either $\rho(x_{k+1}, X_k) < \varepsilon$, in which case we set $X_{k+1} = X_k$, or else $\rho(x_{k+1}, X_k) > 0$. In the latter case we take X_{k+1} to be the span of $X_k \cup \{x_{k+1}\}$, which is finite-dimensional by the preceding lemma. This completes the inductive construction of the finite-dimensional subspace X_{k+1}. It remains to take $Y = X_n$. □

Lemma 4.1.12. ***(Riesz's lemma)*** *Let Y be a closed located subspace with an inhabited metric complement in a normed space X, and let $0 < \theta < 1$. Then there exists a unit vector $x \in X$ such that $\|x - y\| > \theta$ for each $y \in Y$.*

Proof. Fix $x_0 \in -Y$. Then

$$0 < r = \rho(x_0, Y) < \theta^{-1} r.$$

Choosing $y_0 \in Y$ such that

$$r \leqslant \|x_0 - y_0\| < \theta^{-1} r,$$

let

$$x = \frac{1}{\|x_0 - y_0\|}(x_0 - y_0).$$

Then $\|x\| = 1$. Also, for each $y \in Y$,

$$y_0 + \|x_0 - y_0\| \, y \in Y.$$

Hence

$$\|x_0 - y_0\| \, \|x - y\| = \|x_0 - (y_0 + \|x_0 - y_0\| \, y)\| \geqslant \rho(x_0, Y) = r$$

and therefore

$$\|x - y\| \geqslant \frac{r}{\|x_0 - y_0\|} > \theta.$$

□

Proposition 4.1.13. *A locally totally bounded normed space is finite-dimensional.*

Proof. Let B be the unit ball of a locally totally bounded normed space X. Then B is located in X. By Proposition 2.2.18, B is locally totally bounded; being bounded, it is therefore totally bounded. Let $\{x_1, \ldots, x_n\}$ be a 1/4-approximation to B, and, using Lemma 4.1.11, construct a finite-dimensional subspace Y of X such that $\rho\,(x_i, Y) < 1/4$ for each i. Let ξ be any point of X, and suppose that $\rho\,(\xi, Y) > 0$. Then, by Riesz's lemma, there exists a unit vector $x \in X$ such that $\|x - y\| > 1/2$ for all $y \in Y$. But this is absurd: for since $x \in B$, there exists i such that $\|x - x_i\| < 1/4$ and therefore $\rho(x, Y) < 1/2$. We conclude that $\rho\,(\xi, Y) = 0$; so ξ is in the closure of Y. But Y, being finite-dimensional and therefore locally compact (Proposition 4.1.6), is closed in X. Hence $\xi \in Y$. Since $\xi \in X$ is arbitrary, we see that $X = Y$. □

Corollary 4.1.14. *A located subspace of a finite-dimensional normed space is finite-dimensional.*

Proof. Let Y be a located subspace of a finite-dimensional normed space X. Since, by Proposition 4.1.6, X is locally compact, we see from Proposition 2.2.18 that Y is locally totally bounded. Hence, by Proposition 4.1.13, Y is finite-dimensional. □

By a *convex combination* of finitely many elements $x_1, \ldots, x_n$ of X we mean a point of the form $\sum_{i=1}^{n} \lambda_i x_i$ where each $\lambda_i \geqslant 0$ and $\sum_{i=1}^{n} \lambda_i = 1$. The *convex hull* of a subset S of X is the closure of the set of all points of X that are convex combinations of points of S. It is straightforward to show that the convex hull of S is the intersection of all convex subsets of X that contain S.

According to Exercise 11 of Chapter 2, a located subset of $\mathbb{R}$ may not have its metric complement located. However, things are different when the set is convex.

Proposition 4.1.15. *If S is a located convex subset of the product normed space $\mathbb{R}^n$ such that $-S$ is inhabited, then $-S$ is located.*

Proof. Consider any x in $\mathbb{R}^n$, and for each $r > 0$ let $\overline{B}_1(x, r)$ be the closed ball with centre x and radius r relative to the product metric ρ. Then $\overline{B}_1(x, r)$ is an n-dimensional cube with centre x and sides of length $2r$. Consider any two real numbers α, β such that $\alpha < \beta$. Let $v_1, \ldots, v_{2^n}$ be the vertices of $\overline{B}_1\left(x, \frac{1}{2}(\alpha + \beta)\right)$. It is left as an exercise to show that there exists $\delta > 0$ such that for all points $w_1, \ldots, w_{2^n}$ with

$$\rho(v_i, w_i) < \delta \quad (1 \leqslant i \leqslant 2^n)\,,$$

the convex hull of $\{w_1, \ldots, w_{2^n}\}$ contains $\overline{B}_1(x, \alpha)$. Either $\rho(v_i, S) < \delta$ for all i, or $\rho(v_i, S) > 0$ for some i. In the first case, for each i in $\{1, \ldots, 2^n\}$ choose w_i in S such that $\rho(v_i, w_i) < \delta$; then S contains the convex hull of $\{w_i, \ldots, w_{2^n}\}$ and therefore contains $\overline{B}_1(x, \alpha)$. Thus $\rho(x, y) \geqslant \alpha$ for all y in $-S$. On the other hand, if for some i we have $\rho(v_i, S) > 0$, then $v_i \in -S$ and $\rho(x, v_i) \leqslant \frac{1}{2}(\alpha + \beta) < \beta$. Thus $\rho(x, S)$ exists, by the constructive least-upper-bound principle. □

Although we do not develop integration theory in this book, it is worth pointing out here that for convex subsets of $\mathbb{R}^n$ there is a close link between locatedness and Lebesgue measurability: a convex subset of $\mathbb{R}^n$ with inhabited interior is located if and only if it is Lebesgue measurable [17]. Informally, this result shows that a convex subset of $\mathbb{R}^n$ with inhabited interior can be located if and only if its size can be calculated.

4.2 Best Approximation

Let V be an inhabited subspace of a metric space X, and let a, b be elements of X, V respectively. We say that b is a *best approximation*, or *closest point*, to a in V if $\rho(a, b) \leqslant \rho(a, v)$ for each $v \in V$. In that case, $\rho(a, V)$ exists and equals $\rho(a, b)$. We call V *proximinal* in X if each element of X has a best approximation in V, in which case V is located.

The fundamental theorem of classical approximation theory says that a finite-dimensional subspace of a normed space is proximinal. Constructively, this theorem implies LLPO (see Exercise 1). However, by introducing the idea of at most one object even without knowing in advance that there exists one, we can produce a good constructive version of the fundamental theorem.

Let X, V, and a be as in the first paragraph of this section. We say that a has *at most one best approximation* in V if for all distinct points v, v' in V, there exists $x \in V$ such that

$$\max\{\rho(a, v), \rho(a, v')\} > \rho(a, x).$$

We call V *quasiproximinal* if each point of X with at most one best approximation in V actually has a (perforce strongly unique) best approximation in V. Our destination in this section is the following constructive *fundamental theorem of approximation theory.*

Theorem 4.2.1. *Every finite-dimensional subspace of a real normed space is quasiproximinal.*

We defer the proof until we have prepared the pathway with some preliminary results, one of which has a proof that uses the λ-technique discussed in Chapter 3.

Lemma 4.2.2. *Let $\{e_1, \ldots, e_n\}$ be a basis for an n-dimensional subspace V of a normed space X, let $1 \leqslant m < n$, and let W be the subspace of X with basis $\{e_1, \ldots, e_m\}$. Then the span of $\{e_{m+1}, \ldots, e_n\}$ is an $(n-m)$-dimensional subspace of the quotient space X/W.*

Proof. The proof is relegated to Exercise 4. □

Lemma 4.2.3. *Let x, e be elements of a real normed space X with $e \neq 0$, and for each $\delta > \rho(x, \mathbb{R}e)$ write*

$$S_\delta = \{t \in \mathbb{R} : \|x - te\| \leqslant \delta\}.$$

If S_δ is compact, then it is a proper compact interval $[m, M]$ in $\mathbb{R}$. Moreover,

$$\|x - me\| = \delta = \|x - Me\|. \tag{4.4}$$

Proof. Suppose that S_δ is compact, with infimum m and supremum M. By Corollary 2.2.14, there exists δ' such that $\rho(x, \mathbb{R}e) < \delta' < \delta$ and $S_{\delta'}$ is compact. Let m' and M' denote the infimum and supremum, respectively, of $S_{\delta'}$. The uniform continuity of the map $t \rightsquigarrow \|x - te\|$ on $\mathbb{R}$ ensures that (4.4) and

$$\|x - m'e\| = \delta' = \|x - M'e\|$$

hold. Since $S_{\delta'} \subset S_\delta$ and $\delta' < \delta$, it follows that $m < m' \leqslant M' < M$ and therefore $m < M$. Now consider any $t \in [m, M]$. Writing

$$\tau = \frac{t - m}{M - m} \geqslant 0,$$

we have $t = (1 - \tau)m + \tau M$. Hence

$$\|x - te\| \leqslant (1 - \tau)\|x - me\| + \tau\|x - Me\| = (1 - \tau)\delta + \tau\delta = \delta$$

and so $t \in S_\delta$. Thus $S_\delta = [m, M]$. □

Lemma 4.2.4. *Let x, e be elements of a real normed space X with $e \neq 0$, and let $d \geqslant 0$. Suppose that*

$$\max\{\|x - te\|, \|x - t'e\|\} > d$$

whenever t, t' are distinct real numbers. Then there exists $\tau \in \mathbb{R}$ such that $\|x - \tau e\| > d$ entails $\rho(x, \mathbb{R}e) > d$.

Proof. Without loss of generality, we may assume that $\rho(x, \mathbb{R}e) < d + 1$. With S_δ as in Lemma 4.2.3, we construct an increasing binary sequence $(\lambda_n)_{n\geqslant 1}$, sequences $(a_n)_{n\geqslant 1}, (b_n)_{n\geqslant 1}$ of real numbers, and a decreasing sequence $(\delta_n)_{n\geqslant 1}$ of positive numbers with the following properties:

- ▷ S_{δ_n} is the proper compact interval $[a_n, b_n]$;
- ▷ if $\lambda_n = 0$, then $\rho(x, \mathbb{R}e) < \delta_n < d + 1/n$, and if $n \geqslant 2$,

$$0 < b_n - a_n \leqslant \frac{2}{3}(b_{n-1} - a_{n-1}); \tag{4.5}$$

▷ if $\lambda_n = 1$, then $\rho(x, \mathbb{R}e) > d$, $a_n = a_{n-1},\ b_n = b_{n-1}$, and $\delta_n = \delta_{n-1}$.

To begin the construction, set $\lambda_1 = 0$ and, using Corollary 2.2.14, choose $\delta_1 > 0$ such that $\rho(x, \mathbb{R}e) < \delta_1 < d+1$ and S_{δ_1} is compact. By Lemma 4.2.3, S_{δ_1} is a proper compact interval $[a_1, b_1]$. Now suppose we have constructed $\lambda_{n-1}, a_{n-1}, b_{n-1}$, and δ_{n-1} with the applicable properties. If $\lambda_{n-1} = 1$, set

$$\lambda_n = 1,\ a_n = a_{n-1},\ b_n = b_{n-1},\ \text{and } \delta_n = \delta_{n-1}. \tag{4.6}$$

If $\lambda_{n-1} = 0$, write

$$r = \max\left\{\left\|x - \left(\frac{1}{3}a_{n-1} + \frac{2}{3}b_{n-1}\right)e\right\|, \left\|x - \left(\frac{2}{3}a_{n-1} + \frac{1}{3}b_{n-1}\right)e\right\|\right\}.$$

By our hypotheses, $r > d$; so either $\rho(x, \mathbb{R}e) > d$ or else $\rho(x, \mathbb{R}e) < \min\{r, d + 1/n\}$. In the first case we define our numbers as at (4.6). In the second case we set $\lambda_n = 0$ and choose δ_n such that

$$\rho(x, \mathbb{R}e) < \delta_n < \min\left\{\delta_{n-1}, r, d + \frac{1}{n}\right\}$$

and S_{δ_n} is compact; then by Lemma 4.2.3, S_{δ_n} is a proper compact interval $[a_n, b_n]$ contained in $S_{\delta_{n-1}}$. Since $r > \delta_n$, either

$$\left\|x - \left(\frac{1}{3}a_{n-1} + \frac{2}{3}b_{n-1}\right)e\right\| > \delta_n$$

or

$$\left\|x - \left(\frac{2}{3}a_{n-1} + \frac{1}{3}b_{n-1}\right)e\right\| > \delta_n.$$

Taking, for example, the first case, and using the convexity of S_{δ_n}, we see that either

$$S_{\delta_n} \subset \left[a_{n-1}, \frac{1}{3}a_{n-1} + \frac{2}{3}b_{n-1}\right)$$

or

$$S_{\delta_n} \subset \left(\frac{1}{3}a_{n-1} + \frac{2}{3}b_{n-1}, b_{n-1}\right].$$

Hence (4.5) holds. This completes our inductive construction.

We see from (4.5) that for $n \geqslant 2$, if $\lambda_n = 0$, then

$$0 < b_n - a_n \leqslant \left(\frac{2}{3}\right)^{n-1}(b_1 - a_1)$$

and therefore

$$0 \leqslant a_{n+1} - a_n \leqslant \left(\frac{2}{3}\right)^{n-1}(b_1 - a_1). \tag{4.7}$$

Clearly, this last inequality also holds if $\lambda_n = 1$. It follows that $(a_n)_{n\geqslant 1}$ is a Cauchy sequence and therefore converges to a limit τ in $\mathbb{R}$. Supposing that $\|x - \tau e\| > d$, compute a positive integer N such that $\|x - a_N e\| > d + 1/N$. If $\lambda_N = 0$, then

$$\|x - a_N e\| = \delta_N < d + \frac{1}{N},$$

a contradiction. We conclude that $\lambda_N = 1$ and hence that $\rho(x, \mathbb{R}e) > d$. $\square$

We now have the proof of Theorem 4.2.1.

Proof. Let V be a finite-dimensional subspace of a normed space X over $\mathbb{R}$, and let a be a point of X with at most one best approximation in V. If V has dimension 0, then there is nothing to prove. If $\dim(V) = 1$, then $V = \mathbb{R}e$ for some $e \neq 0$ in V, and we can apply Lemma 4.2.4 with $x = a$ and $d = \rho(a, V)$ to construct $\tau \in \mathbb{R}$ such that $\|a - \tau e\| = \rho(a, V)$.

Now let n be a positive integer, and suppose we have proved the desired result for all n-dimensional subspaces of real normed spaces. Consider the case where V has a basis $\{e_1, \dots, e_{n+1}\}$, and let $Y = \mathbb{R}e_{n+1}$. By Lemma 4.2.2, V/Y is an n-dimensional subspace of X/Y with basis $\{e_1, \dots, e_n\}$; moreover, for each $x \in X$ we have

$$\rho(x, V) = \inf\left\{\|x - v\|_{X/Y} : v \in V\right\}. \tag{4.8}$$

Next, note that

(*) *for each $v \in V$ there exists $\alpha \in \mathbb{R}$ such that if $\|a - v - \alpha e_{n+1}\| > \rho(a, V)$, then*

$$\|a - v\|_{X/Y} = \rho(a - v, \mathbb{R}e_{n+1}) > \rho(a, V).$$

For if t, t' are distinct real numbers, then

$$\|(v + te_{n+1}) - (v + t'e_{n+1})\| = |t - t'|\,\|e_{n+1}\| \neq 0$$

and therefore, by our hypotheses,

$$\max\{\|a - v - te_{n+1}\|, \|a - v - t'e_{n+1}\|\} > \rho(a, V).$$

So we can apply Lemma 4.2.4 with $x = a - v$, $e = e_{n+1}$, and $d = \rho(a, V)$ to compute the desired α.

With v, α as above, now let v' be a point of V distinct from v, and compute $\alpha' \in \mathbb{R}$ such that if $\|a - v' - \alpha' e_{n+1}\| > \rho(a, V)$, then

$$\|a - v'\|_{X/Y} = \rho(a - v', \mathbb{R}e_{n+1}) > \rho(a, V).$$

We have

$$\begin{aligned}\|(v+\alpha e_{n+1})-(v'+\alpha' e_{n+1})\| &= \|(v-v')+(\alpha-\alpha')\,e_{n+1}\| \\ &\geqslant \|v-v'\|_{X/Y}>0,\end{aligned}$$

so, by our hypotheses,

$$\max\{\|a-v-\alpha e_{n+1}\|,\|a-v'-\alpha' e_{n+1}\|\}>\rho(a,V).$$

It follows from (4.8) and the defining properties of α,α' that

$$\max\left\{\|a-v\|_{X/Y},\|a-v'\|_{X/Y}\right\}>\rho(a,V)=\inf\left\{\|a-v\|_{X/Y}:v\in V\right\}.$$

We have now shown that a has at most one best approximation in the n-dimensional subspace V/Y of X/Y. By our induction hypothesis and (4.8), there exists $v_0\in V$ such that

$$\rho(a-v_0,\mathbb{R}e_{n+1})=\|a-v_0\|_{X/Y}=\inf\left\{\|a-v\|_{X/Y}:v\in V\right\}=\rho(a,V).$$

Applying (*) once more, we compute $t_0\in\mathbb{R}$ such that

$$\|a-v_0-t_0e_{n+1}\|=\rho(a,V).$$

Hence $v_0+t_0e_{n+1}$ is the best approximation to a with respect to the original norm on X, and our inductive proof is complete. □

Theorem 4.2.1 is classically equivalent to its classical counterpart. For if V is a finite-dimensional subspace of the real normed space X and we suppose that some $a\in X$ has no best approximation in V, then a clearly has at most one best approximation in X. Therefore, by Theorem 4.2.1, it has one, which is a contradiction.

For applications of Theorem 4.2.1, see [15].

4.3 Hilbert Spaces

As we did for metric and normed spaces, we assume that the reader is familiar with the definitions and elementary properties of an *inner product* and an *inner product space.* In particular, we are not going to prove that

$$\|x\|=\sqrt{\langle x,x\rangle}$$

defines the norm associated with a given inner product on a vector space X over $\mathbb{K}$, and that the *Cauchy–Schwarz inequality*

$$|\langle x,y\rangle|\leqslant\|x\|\,\|y\|$$

and the *parallelogram identity*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

then hold. The latter inequality enables us to prove that an inner product space X is uniformly convex, as follows: If $0 < \delta < 1$, and x, y are unit vectors in X with $\frac{1}{2}\|x+y\| > 1-\delta$, then

$$\|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 - \|x+y\|^2 < 4 - 4(1-\delta)^2,$$

which can be made as small as we please, independently of x and y, by a suitable initial choice of δ.

The simplest example of an inner product space is, of course, the Euclidean space $\mathbb{K}^n$ with the usual inner product given by

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i^*,$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Another example is the space $l_2(\mathbb{K})$, consisting of all sequences $\mathbf{x} = (x_n)_{n \geqslant 1}$ in $\mathbb{K}$ that are *square summable* in the sense that $\sum_{n=1}^{\infty} |x_n|^2$ converges in $\mathbb{R}$; in this case we work with termwise operations and with the inner product defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n=1}^{\infty} x_n y_n^*.$$

It is left to the exercises to prove that this is indeed an inner product.

An inner product space that is complete with respect to its norm is called a *Hilbert space.* Adding completeness to the inner product leads to a structure with powerful geometrical properties such as the following.

Theorem 4.3.1. *Let S be a closed, located subspace of a Hilbert space H. Then for each $x \in H$, there exists a strongly unique element Px of S such that $\|x - Px\| = \rho(x, S)$. Moreover, Px is the strongly unique element y of S such that $\langle x-y, s\rangle = 0$ for all $s \in S$.*

Proof. Fixing $x \in H$, let $d = \rho(x, S)$ and choose a sequence $(s_n)_{n \geqslant 1}$ in S such that $d = \lim_{n\to\infty} \|x - s_n\|$. Using the parallelogram identity and the convexity of S, we have

$$\begin{aligned}
\|s_m - s_n\|^2 &= \|(s_m - x) - (s_n - x)\|^2 \\
&= 2\|s_m - x\|^2 + 2\|s_n - x\|^2 - 4\left\|\frac{s_m + s_n}{2} - x\right\|^2 \\
&\leqslant 2\|s_m - x\|^2 + 2\|s_n - x\|^2 - 4d^2 \\
&= 2\left(\|s_m - x\|^2 - d^2\right) + 2\left(\|s_n - x\|^2 - d^2\right) \\
&\longrightarrow 0 \quad \text{as } m, n \longrightarrow \infty.
\end{aligned}$$

Hence $(s_n)_{n\geqslant 1}$ is a Cauchy sequence. Since H is complete and S is closed, this sequence converges to a limit $Px \in S$. The continuity of the norm on H now gives $\|x - Px\| = d$. Moreover, if $y \in S$ and $y \neq Px$, then, again by the parallelogram identity and convexity, we have

$$\begin{aligned}
0 < \|Px - y\|^2 &= \|(Px - x) - (y - x)\|^2 \\
&= 2\|Px - x\|^2 + 2\|y - x\|^2 - 4\left\|\frac{Px + y}{2} - x\right\|^2 \\
&\leqslant 2\left(\|Px - x\|^2 - d^2\right) + 2\left(\|y - x\|^2 - d^2\right) \\
&= 2\left(\|y - x\|^2 - d^2\right),
\end{aligned}$$

so $\|x - y\| > d$. It follows that Px is the strongly unique closest point to x in S.

Next note that for all $y \in S$ and $\lambda \in K$,

$$\langle x - Px + \lambda y, x - Px + \lambda y\rangle \geqslant d^2 = \langle x - Px, x - Px\rangle$$

and therefore

$$|\lambda|^2 \|y\|^2 + 2\,\mathrm{Re}\left(\lambda^* \langle x - Px, y\rangle\right) \geqslant 0. \tag{4.9}$$

Suppose that $\mathrm{Re}\langle x - Px, y\rangle \neq 0$. By choosing a sufficiently small real number λ with $\lambda\,\mathrm{Re}\langle x - Px, y\rangle < 0$, we can contradict (4.9). It follows that $\mathrm{Re}\langle x - Px, y\rangle = 0$. A similar argument shows that $\mathrm{Im}\langle x - Px, y\rangle = 0$; whence $\langle x - Px, y\rangle = 0$. Finally, if $y \in S$ and $y \neq Px$, then since $y - Px \in S$,

$$\begin{aligned}
0 < \|y - Px\|^2 &= \langle y - Px, y - Px\rangle \\
&= \langle x - Px, y - Px\rangle - \langle x - y, y - Px\rangle = -\langle x - y, y - Px\rangle,
\end{aligned}$$

so $\langle x - y, y - Px\rangle \neq 0$. It follows that Px is the strongly unique element y of S such that $\langle x - y, s\rangle = 0$ for all $s \in S$. □

The mapping $P : H \longrightarrow S$ defined in Theorem 4.3.1 is called the *projection of H onto S*, and for each $x \in H$, the vector Px is the *projection of the vector x onto S*.

The mapping P is linear: for if $\lambda \in \mathbb{K}$ and $x, x' \in H$, then since for all $y \in S$,

$$\langle \lambda x + x' - (\lambda Px + Px'), y\rangle = \lambda\langle x - Px, y\rangle + \langle x' - Px', y\rangle = 0,$$

it follows from the uniqueness of the projection of $\lambda x + x'$ onto S (Theorem 4.3.1) that $P(\lambda x + x') = \lambda Px + Px'$. For each $y \in S$, since $\langle y - y, s\rangle = 0$ for all $s \in S$, we see from the uniqueness of the projection of y into S that $Py = y$; whence P maps H onto S. Since $Px \in S$, we have $P^2 x = P(Px) = Px$, so P is *idempotent*; that is, $P^2 = P$. Also,

$$\begin{aligned}\|x\|^2 = \langle x, x\rangle &= \langle x - Px + Px, x - Px + Px\rangle \\ &= \langle x - Px, x - Px\rangle + \langle x - Px, Px\rangle + \langle Px, Px\rangle \\ &= \|x - Px\|^2 + \|Px\|^2,\end{aligned}$$

since $\langle x - Px, Px\rangle = 0$ (by Theorem 4.3.1). Hence $\|Px\| \leqslant \|x\|$, and so 1 is a bound for P. On the other hand, if S contains a nonzero vector x, then $\|Px\| = \|x\|$; it follows that in this case, P is normed and $\|P\| = 1$.

Two subsets S, T of an inner product space are said to be *orthogonal* if $\langle x, y\rangle = 0$ for all $x \in S$ and $y \in T$; we then write $S \perp T$. The orthogonality relation is symmetric,

$$S \perp T \Longleftrightarrow T \perp S,$$

and any family $(T_i)_{i\in I}$ of subsets of X satisfies

$$S \perp \bigcup_{i\in I} T_i \Longleftrightarrow \forall i \in I\,(S \perp T_i).$$

A *vector x is orthogonal to the subset S* if $\{x\} \perp S$, in which case we write $x \perp S$; two vectors x, y are said to be *orthogonal vectors* if $x \perp \{y\}$, in which case we write $x \perp y$ and we have the following generalisation of *Pythagoras's theorem*:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

We define the *orthogonal complement* of a subset S of X to be the set

$$S^\perp = \{x \in X : x \perp S\},$$

which is easily seen to be a closed linear subspace of X.

If $0 \in S$, then $S \cap S^\perp = \{0\}$. It readily follows that if S is a closed, located subspace of a Hilbert space H, with P the corresponding projection, and if $x \in H$, then the decomposition

$$x = Px + (x - Px)$$

is the unique expression of x as the sum of a vector in S and a vector orthogonal to S. Denoting by I the *identity operator* $x \rightsquigarrow x$ on H, we prove that $\mathrm{I} - P$ is the projection of H onto the subspace $S^\perp$. For each $y \in S^\perp$, the vector $z = x - Px - y$ belongs to $S^\perp$, so

$$\|x - y\|^2 = \|Px + z\|^2 = \|Px\|^2 + \|z\|^2 \geqslant \|Px\|^2 = \|x - (x - Px)\|^2.$$

Since $x - Px \in S^\perp$, we conclude from the uniqueness part of Theorem 4.3.1 that $x - Px$ is the projection of x into $S^\perp$.

We need at this point to clarify what we mean by saying that *the series* $\sum_{i\in I} \lambda_i e_i$ *converges to the sum* x and by writing

$$x = \sum_{i\in I} \lambda_i e_i, \tag{4.10}$$

where the index set I on the right is not necessarily countable: we mean that for each $\varepsilon > 0$ there exists a finitely enumerable subset F of I with the property that if G is a finitely enumerable subset of I that contains F, then

$$\left\| x - \sum_{i \in G} \lambda_i e_i \right\| < \varepsilon.$$

In the case $I = \mathbb{N}^+$, this condition is equivalent to the usual one for convergence of the series $\sum_{n=1}^{\infty} \lambda_n e_n$ to x.

A family $(e_i)_{i \in I}$ of vectors in a Hilbert space H is said to be *orthonormal* if

- ▷ $e_i \perp e_j$ whenever $i \neq j$, and
- ▷ for each i, either $\|e_i\| = 1$ or $e_i = 0$.

Such a family is called an *orthonormal basis* if each vector $x \in H$ can be written uniquely in the form (4.10) with $(\lambda_i)_{i \in I}$ a family of elements of $\mathbb{K}$ such that $\lambda_i = 0$ whenever $e_i = 0$. The scalar λ_i is then called the ith *coordinate* of x relative to the orthonormal basis.

If $(e_i)_{i \in I}$ is a finitely enumerable orthonormal family in H, then it spans a finite-dimensional subspace of H. To see this, we may assume that $\|e_i\| = 1$ for each i. Let $(\lambda_i)_{i \in I}$ be a family of scalars such that $\sum_{i \in I} |\lambda_i| > 0$, and let $x = \sum_{i \in I} \lambda_i e_i$. Choose $j \in I$ such that $\lambda_j \neq 0$. Then

$$\langle x, e_j \rangle = \sum_{i \in I} \lambda_i \langle e_i, e_j \rangle = \lambda_j \langle e_j, e_j \rangle = \lambda_j \neq 0,$$

so $x \neq 0$. Thus the vectors e_i $(i \in I)$ are linearly independent and therefore span a finite-dimensional space.

Classically, using (an equivalent of) the axiom of choice, we can prove that every Hilbert space has an orthonormal basis of unit vectors. Constructively we avoid the axiom of choice by adding separability to the hypotheses on H, by relaxing the requirements to allow basis vectors to be 0, and by using the *Gram–Schmidt orthogonalisation process* embodied in the proof of our next result.

Proposition 4.3.2. *Every separable Hilbert space has a countable orthonormal basis. If $(e_n)_{n \geqslant 1}$ is such a basis, relative to which x has coordinates $\alpha_1, \alpha_2, \ldots$ and y has coordinates $\beta_1, \beta_2, \ldots$, then $\alpha_n = \langle x, e_n \rangle$ for each n, $\langle x, y \rangle = \sum_{n=1}^{\infty} \alpha_n \beta_n^*$, and $\|x\|^2 = \sum_{n=1}^{\infty} |\alpha_n|^2$.*

Proof. Let $(a_n)_{n \geqslant 1}$ be a dense sequence in H. The idea of the proof is to construct inductively the orthonormal sequence $(e_n)_{n \geqslant 1}$ to ensure that at stage n there exists

z_n in the subspace S_n of H spanned by $\{e_1, \ldots, e_n\}$ such that $\|a_n - z_n\| < 1/n$. For convenience, set $e_0 = 0$ and $S_0 = \{0\}$. Assume that e_n has already been constructed. Being finite-dimensional, S_n is closed and located, so the projection P_n of H onto S_n exists. Either $\|a_{n+1} - P_n a_{n+1}\| < 1/(n+1)$ or else $a_{n+1} \neq P_n a_{n+1}$. In the first case, set $e_{n+1} = 0$. In the second, define a unit vector by

$$e_{n+1} = \frac{1}{\|a_{n+1} - P_n a_{n+1}\|} (a_{n+1} - P_n a_{n+1}).$$

Note that e_{n+1} is orthogonal to $e_0, \ldots, e_n$ since (by Theorem 4.3.1) $a_{n+1} - P_n a_{n+1}$ is orthogonal to S_n, and that

$$a_{n+1} = P_n a_{n+1} + \|a_{n+1} - P_n a_{n+1}\| e_{n+1}$$

is in the subspace of H spanned by $\{e_1, \ldots, e_{n+1}\}$. This completes the induction.

Now consider the vectors $x, y \in H$. For each positive integer k we have unique representations

$$P_k x = \sum_{n=1}^{k} \alpha_n e_n$$

and

$$P_k y = \sum_{n=1}^{k} \beta_n e_n,$$

with $\alpha_n = \beta_n = 0$ whenever $e_n = 0$. For $n \leqslant k$,

$$\alpha_n = \sum_{i=1}^{k} \alpha_i \langle e_i, e_n \rangle = \langle P_k x, e_n \rangle = \langle x, e_n \rangle - \langle x - P_k x, e_n \rangle = \langle x, e_n \rangle .$$

Likewise, $\beta_n = \langle y, e_n \rangle$. Thus α_n and β_n do not depend on any $k \neq n$. Now,

$$\langle P_k x, P_k y \rangle = \sum_{n=1}^{k} \alpha_n \beta_n^*.$$

Since the sequence $(a_n)_{n \geqslant 1}$ is dense in X and $\rho(a_k, S_k) < 1/k$, we see that both $\rho(x, S_k)$ and $\rho(y, S_k)$ approach 0 as $k \longrightarrow \infty$; whence $P_k x \longrightarrow x$ and $P_k y \longrightarrow y$ as $k \longrightarrow \infty$. Thus

$$\langle x, y \rangle = \lim_{k \to \infty} \langle P_k x, P_k y \rangle = \sum_{n=1}^{\infty} \alpha_n \beta_n^*.$$

Finally, if we take $y = x$, we obtain $\|x\|^2 = \sum_{n=1}^{\infty} |\alpha_n|^2$. □

An elementary lemma will enable us to characterise dimensionality in terms of orthonormal bases.

Lemma 4.3.3. *Let S be a finite-dimensional subspace of a separable Hilbert space H, and let $(e_n)_{n\geqslant 1}$ be an orthonormal basis of H. Then there exists a positive integer N such that $\rho(e_n, S) > 0$ whenever $n \geqslant N$ and $\|e_n\| = 1$.*

Proof. By Corollary 4.1.7, there exists a finite 1/2-approximation $\{x_1, \ldots, x_m\}$ to the unit ball of H. Since, by the preceding proposition, the series $\sum_{n=1}^{\infty} |\langle x_k, e_n\rangle|^2$ converges, $|\langle x_k, e_n\rangle| \longrightarrow 0$ as $n \longrightarrow \infty$; so we can compute N such that

$$|\langle x_k, e_n\rangle| < \frac{3}{8} \quad (1 \leqslant k \leqslant m;\ n \geqslant N).$$

Let P be the projection of H on the finite-dimensional, hence closed and located, subspace S. Consider any $n \geqslant N$ such that $\|e_n\| = 1$. Choosing k $(1 \leqslant k \leqslant m)$ such that $\|Pe_n - x_k\| < 1/2$, we have

$$\|e_n - x_k\|^2 = 1 + \|x_k\|^2 - 2\,\mathrm{Re}\,\langle x_k, e_n\rangle \geqslant 1 - 2\,|\langle x_k, e_n\rangle| > \frac{1}{4},$$

so

$$\rho(e_n, S) = \|e_n - Pe_n\| \geqslant \|e_n - x_k\| - \|Pe_n - x_k\| > 0,$$

as we required. □

Proposition 4.3.4. *Let $(e_n)_{n\geqslant 1}$ be an orthonormal basis of a separable Hilbert space H. Then H is finite-dimensional if and only if $e_n = 0$ for all sufficiently large n.*

Proof. If H is finite-dimensional, then taking $S = H$ in Lemma 4.3.3, we obtain N such that $e_n = 0$ for all $n \geqslant N$. If, conversely, such N exists, then H is the finite-dimensional space of all linear combinations of those vectors $e_1, \ldots, e_N$ that are nonzero. □

Let X be a linear space with an inequality compatible with its linear structure. We say that X is *infinite-dimensional* if the complement of each finite-dimensional subspace of X is inhabited. It follows from Bishop's lemma (Proposition 3.1.1) that a normed space X is infinite-dimensional if and only if for each finite-dimensional subspace S of X there exists $x \in X$ such that $\rho(x, S) > 0$.

Proposition 4.3.5. *Let $(e_n)_{n\geqslant 1}$ be an orthonormal basis of a separable Hilbert space H. Then H is infinite-dimensional if and only if $e_n \neq 0$ for infinitely many n.*

Proof. Suppose that H is infinite-dimensional, and consider any positive integer N. Let S be the finite-dimensional subspace of H consisting of all linear combinations of $e_1, \dots, e_N$, and let P be the projection of H onto S. There exists $x \in H$ such that $\rho(x, S) > 0$ and therefore, by Proposition 4.3.2,

$$\sum_{n=N+1}^{\infty} |\langle x, e_n \rangle|^2 = \|x - Px\|^2 = \rho(x, S)^2 > 0.$$

Hence there exists $n > N$ such that $\langle x, e_n \rangle \neq 0$ and therefore, by the Cauchy–Schwarz inequality, $e_n \neq 0$.

Conversely, if there exists a strictly increasing sequence $(n_k)_{k \geqslant 1}$ of positive integers such that $e_{n_k} \neq 0$ for each k, consider any finite-dimensional subspace S of H. Choose N as in Lemma 4.3.3. Then choose k such that $n_k > N$. Since $\|e_{n_k}\| = 1$, we must have $\rho(e_{n_k}, S) > 0$. □

We now introduce a construction of new Hilbert spaces from old. Let H_1, H_2 be Hilbert spaces, and for all $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ in the product vector space $H_1 \times H_2$ define

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle .$$

This defines an inner product with respect to which $H_1 \times H_2$ is a Hilbert space, called the *direct sum* of H_1 and H_2, and denoted by $H_1 \oplus H_2$. This construction helps us to prove the most general case of our next theorem.

We say that a linear functional u on a Hilbert space H is *represented* by a vector $a \in H$ if $u(x) = \langle x, a \rangle$ for all $x \in H$. The following *Riesz representation theorem* tells us that a functional is representable by a vector in this way if and only if it is normed.

Theorem 4.3.6. *A bounded linear functional u on a Hilbert space H is normed if and only if there exists $a \in H$ such that*

$$u(x) = \langle x, a \rangle \quad (x \in H). \tag{4.11}$$

In that case, a is strongly unique: if $y \in H$ and $y \neq a$, then there exists $x \in H$ such that $u(x) \neq \langle x, y \rangle$. Moreover, $\|u\| = \|a\|$.

Proof. Suppose first that there exists a vector a with property (4.11). Then, by the Cauchy–Schwarz inequality, $|u(x)| \leqslant \|a\| \|x\|$ for each $x \in H$. On the other hand, given $\varepsilon > 0$, we have either $\|a\| > 0$ or $\|a\| < \varepsilon$. In the first case,

$$u\left(\frac{1}{\|a\|} a\right) = \|a\| > \|a\| - \varepsilon.$$

In the second, $u(0) = 0 > \|a\| - \varepsilon$. Since $\varepsilon > 0$ is arbitrary, it follows that $\|u\|$ exists and equals $\|a\|$.

Suppose, conversely, that u is normed. To begin with, take $\|u\| > 0$. Then $\ker(u)$ is located, by Proposition 2.3.6. Let P be the projection of H on $\ker(u)$, choose $y \in H$ with $u(y) > 0$, and define

$$x_0 = \frac{1}{u(y)} (y - Py) .$$

Then $u(x_0) = 1$, so $x - u(x)x_0 \in \ker(u)$. Since $x_0 \perp \ker(u)$, we have

$$0 = \langle x - u(x)x_0, x_0 \rangle = \langle x, x_0 \rangle - u(x) \|x_0\|^2 ,$$

from which it follows that $u(x) = \langle x, a \rangle$ with

$$a = \frac{1}{\|x_0\|^2} x_0.$$

It remains to remove the condition that $\|u\| > 0$. In doing so, we need to bear in mind that we have not ruled out the possibility that $H = \{0\}$. To deal with this, we consider the direct sum $H \oplus \mathbb{K}$, on which we define a bounded linear functional v by

$$v(x, \zeta) = u(x) + \zeta.$$

We first observe that u is represented by the vector a if and only if v is represented by the vector $(a, 1)$. Since $\|v\| > 0$, it follows from the first part of the proof that it will suffice to prove that v is normed. To this end, let $0 < \alpha < \beta$ and set $\varepsilon = \frac{1}{2}(\alpha + \beta)$. If $\|u\| > 0$, then, by the first part of the proof, u is represented by a unique vector $a \in H$; whence v is represented by $(a, 1)$ and is therefore normed. So we may assume that $\|u\| < \varepsilon$. Then either $\|u\| + 1 > \alpha + \varepsilon$ or $\|u\| + 1 < \beta$. In the first case, $v(0, 1) = 1 > \alpha$ and $\|(0, 1)\| = 1$. In the second case, for each (x, ζ) with $\|(x, \zeta)\| \leqslant 1$ we have

$$|v(x, \zeta)| \leqslant |u(x)| + |\zeta| \leqslant \|u\| + 1 < \beta.$$

It now follows from the least-upper-bound principle that $\|v\|$ exists. $\square$

The second part of the foregoing proof contains an argument showing that the sum of two normed linear functionals on a Hilbert space is normed. Note that in general, the sum of two normed linear functionals on an arbitrary nontrivial normed space need not be normed.

By an *operator* on a Hilbert space H we mean a linear mapping of H into itself. The set of bounded operators on H is denoted by $\mathcal{B}(H)$.

Classically, the Riesz representation theorem enables us to prove, for a given element T of $\mathcal{B}(H)$, the existence of the adjoint T^*, which has the defining property

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \quad (x, y \in H) . \tag{4.12}$$

Indeed, given $y \in H$, we apply the Riesz representation theorem to the linear functional $x \rightsquigarrow \langle Tx, y \rangle$ to obtain a unique vector T^*y with the desired property. The constructive problem with this argument is that in order to apply Theorem 4.3.6, we require that the functional $x \rightsquigarrow \langle Tx, y \rangle$ be normed, which is something we cannot guarantee. In fact, as the following Brouwerian example shows, we cannot prove that a general normed operator on H has an adjoint.

Let H be a complex Hilbert space with an orthonormal basis $(e_n)_{n \geqslant 1}$ of unit vectors, and let $(a_n)_{n \geqslant 1}$ be a binary sequence with at most one term equal to 1. Observe that for each $x \in H$ the series $\sum_{n=1}^{\infty} a_n \langle x, e_n \rangle$ converges absolutely. For, given $\varepsilon > 0$ and choosing N such that $\sum_{n=N+1}^{\infty} |\langle x, e_n \rangle|^2 < \varepsilon^2$, for each $k > N$ we have

$$\sum_{n=N+1}^{k} |a_n \langle x, e_n \rangle| = \sum_{n=N+1}^{k} a_n |\langle x, e_n \rangle| \leqslant \max_{N+1 \leqslant n \leqslant k} |\langle x, e_n \rangle| < \varepsilon.$$

Hence the partial sums of $\sum_{n=1}^{\infty} a_n |\langle x, e_n \rangle|$ form a Cauchy sequence, and so the series $\sum_{n=1}^{\infty} a_n \langle x, e_n \rangle$ converges in $\mathbb{C}$. Thus

$$Tx = \left(\sum_{n=1}^{\infty} a_n \langle x, e_n \rangle \right) e_1$$

defines a mapping—clearly an operator—from H to itself. It is left as an exercise to show that T is normed. Suppose that T^* exists. Then either $T^*e_1 \neq 0$ or else $\|T^*e_1\| < 1$. In the first case we can find N such that $\langle T^*e_1, e_N \rangle \neq 0$. Suppose that $a_N = 0$. If there exists $n \neq N$ such that $a_n = 1$, then for all $x, y \in H$,

$$\langle Tx, y \rangle = \langle \langle x, e_n \rangle e_1, y \rangle = \langle x, \langle y, e_1 \rangle e_n \rangle$$

and therefore

$$T^*y = \langle y, e_1 \rangle e_n. \tag{4.13}$$

Hence $\langle T^*e_1, e_N \rangle = 0$, a contradiction. Thus $a_n = 0$ for all $n \neq N$ and therefore for all n; whence $T^* = 0$, which is impossible since $T^*e_1 \neq 0$. We conclude that $a_N = 1$. In the case $\|T^*e_1\| < 1$, suppose that $a_n = 1$. Then (4.13) holds, so $T^*e_1 = e_n$ and therefore $\|T^*e_1\| = 1$, a contradiction. It follows that in this case we have $a_n = 0$ for all n.

In view of this Brouwerian example, for any Hilbert space H and any not necessarily bounded operator T on H, we define the *adjoint* T^*, if it exists, by the equation (4.12); in which case we refer to T as *jointed*. It is then straightforward to show that T^* is an operator; that the adjoint of T^* is T; and that any bound for T is one for T^*, and vice versa. Moreover, if S, T are jointed operators, then for each $\lambda \in \mathbb{K}$, so are $\lambda S + T$ and ST, and

$$(\lambda S+T)^* = \lambda^* S^* + T^*,$$
$$(ST)^* = T^* S^*.$$

For example, we have

$$\langle STx, y\rangle = \langle Tx, S^*y\rangle = \langle x, T^*(S^*y)\rangle = \langle x, T^*S^*y\rangle .$$

An operator T on H is called *selfadjoint*, or *Hermitian*, if T^* exists and equals T. The identity operator I is trivially selfadjoint. More generally, if P is the projection of H onto a closed, located subspace, then for all $x, y \in H$ we have

$$\begin{aligned}\langle Px, y\rangle &= \langle Px, Py\rangle + \langle Px, y-Py\rangle = \langle Px, Py\rangle \\ &= \langle Px, Py\rangle + \langle x-Px, Py\rangle = \langle x, Py\rangle ,\end{aligned}$$

so P is selfadjoint.

Conversely, if P is any bounded, idempotent, selfadjoint operator on H, then P is a projection. To see this, let

$$V = \{y \in H : Py = y\}.$$

It is clear that $V$is a linear subspace of H. If $x \in H$, then since $P^2x = Px$, we have $Px \in V$. On the other hand, for each $y \in V$ we have

$$\langle x-Px, y\rangle = \langle x, y\rangle - \langle Px, y\rangle = \langle x, y\rangle - \langle x, Py\rangle = 0,$$

since $Py = y$. Hence $x - Px$ is orthogonal to V, and therefore (since $Px - y$ is in V)

$$\begin{aligned}\|x-y\|^2 &= \|x-Px+Px-y\|^2 \\ &= \|x-Px\|^2 + \|Px-y\|^2 \geqslant \|x-Px\|^2 .\end{aligned}$$

We now see that Px is a closest point to x in V. Hence V is located in H. Finally, the continuity of the bounded operator P ensures that V is closed; so P is the projection of H onto V.

We end the chapter with an application of the Riesz representation theorem and a corollary, both due to Ishihara [58]. For this we need to know that a linear mapping T between normed spaces X, Y is defined to be *compact* if $T(\overline{B}_X(0,1))$ is a totally bounded subset of Y; in that case, the norm of T exists, by Corollary 2.2.7. Every bounded linear mapping on a finite-dimensional normed space is compact.

Proposition 4.3.7. *Let T be a bounded linear mapping of a Hilbert space H into $\mathbb{C}^n$, and for $1 \leqslant i \leqslant n$ let $P_i : \mathbb{C}^n \longrightarrow \mathbb{C}$ be the ith projection mapping, defined by*

$$P_i(z_1, z_2, \ldots, z_n) = z_i.$$

Then T is compact if and only if $P_i \circ T$ is normed for each i.

Proof. If T is compact, then since the projection P_i is uniformly continuous, the set $P_i \circ T(\overline{B}(0,1))$ is totally bounded, so $\|P_i \circ T\|$ exists. Suppose, conversely, that $P_i \circ T$ is normed for each i. By the Riesz representation theorem, for each i there exists $a_i \in H$ such that

$$P_i \circ T(x) = \langle x, a_i \rangle \quad (x \in H).$$

Using Lemma 4.1.11, construct a finite-dimensional subspace X of H such that $\rho(a_i, X) < \varepsilon/2n$ for each i. Let P be the projection of H onto X. Since the restriction of T to X is a compact linear mapping, there exist $x_1, x_2, \ldots, x_m$ in the unit ball B_X of X such that $\{Tx_1, Tx_2, \ldots, Tx_m\}$ is an $\varepsilon/2$-approximation to $T(B_X)$. Consider any $x \in H$ with $\|x\| \leqslant 1$. Working with the product norm on $\mathbb{C}^n$, we have

$$\begin{aligned}
\|T(x - Px)\| &= \sum_{i=1}^{n} |P_i \circ T(x - Px)| \\
&= \sum_{i=1}^{n} \langle x - Px, a_i \rangle \\
&= \sum_{i=1}^{n} \langle x, (\mathrm{I} - P)\, a_i \rangle \\
&\leqslant \sum_{i=1}^{n} \|x\| \, \|a_i - Pa_i\| < \sum_{i=1}^{n} \frac{\varepsilon}{2n} = \frac{\varepsilon}{2}.
\end{aligned}$$

Choosing j such that $\|TPx - Tx_j\| < \varepsilon/2$, we now obtain

$$\|Tx - Tx_j\| \leqslant \|T(x - Px)\| + \|TPx - Tx_j\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus $\{Tx_1, Tx_2, \ldots, Tx_m\}$ is an ε-approximation to $T(B_H)$. Since $\varepsilon > 0$ is arbitrary, we conclude that T is a compact linear mapping on H. $\square$

Corollary 4.3.8. *The sum of two compact operators on a Hilbert space is compact.*

Proof. Let S and T be compact operators on a Hilbert space H, and let $\varepsilon > 0$. Since $S(B_H)$ and $T(B_H)$ are totally bounded, it readily follows from Lemma 4.1.11 that there exists a finite-dimensional subspace X of H, with associated projection P, such that

$$\|(S + T)x - (PS + PT)x\| < \varepsilon \qquad (x \in B_H). \tag{4.14}$$

We may assume that X is nontrivial and so has a basis $\{e_1, \ldots, e_n\}$. Writing

$$PSx = \sum_{i=1}^{n} p_i(x) e_i, \quad PTx = \sum_{i=1}^{n} q_i(x) e_i,$$

we see from Proposition 4.3.7 that for each i, the linear functionals p_i, q_i on H are normed; whence $p_i + q_i$ is normed (see the remark following the proof of the Riesz representation theorem). Since

$$(PS + PT)(x) = \sum_{i=1}^{n} (p_i + q_i)(x) e_i,$$

we see from Proposition 4.3.7 that $PS + PT$ is compact. In view of (4.14), any ε-approximation to $(PS + PT)(B_H)$ is a 2ε-approximation to $(S+T)(B_H)$. Hence $S + T$ is compact. □

Exercises

1. Let X be the space $\mathbb{R}^2$ with norm $\|(x, y)\| = \max\{|x|, |y|\}$. Show that if the 1-dimensional subspace $\mathbb{R}(\cos\theta, \sin\theta)$ is proximinal in X for each $\theta \in \mathbb{R}$, then LLPO holds.

2. Let V be a nonzero linear subspace of a normed linear space X such that each $x \in X$ has at most one closest point in V. Need V be located?

3. Prove that if every pair of nonzero vectors in $\mathbb{R}^2$ generates a finite-dimensional space, then LPO holds.

4. Prove Lemma 4.2.2.

5. Prove that a uniformly convex linear subspace of a normed space is proximinal.

6. A normed space X is said to be *compactly generated* if there exists a compact set $K \subset X$ such that each point of X is a linear combination of finitely many points of K. Let X be a compactly generated Banach space, and Y a finite-dimensional subspace of X. Prove that either $X = Y$ or else the metric complement of Y is inhabited. (*Hint*: Use the λ-technique.)

7. Let Y be a locally compact subspace of a complete metric space X such that each $x \in X$ has a unique best approximation Px in Y. Prove that the mapping P is sequentially continuous on X. (*Hint:* First use Ishihara's tricks.)

8. Let T be a bounded linear mapping of a normed space X onto a finite-dimensional Banach space Y. Prove that there exists $r > 0$ such that $B_Y(0, r) \subset T(B_X(0, 1))$.

9. Let T be a bounded linear mapping of a normed space X onto a finite-dimensional Banach space Y. Prove that T is a compact linear mapping if and only if $\ker(T)$ is located in X.

10. Let X be an inner product space. Prove that $\|x\| = \sqrt{\langle x, x\rangle}$ defines a norm on X and that the Cauchy–Schwarz inequality $|\langle x, y\rangle| \leqslant \|x\| \, \|y\|$ holds for all $x, y \in X$.

11. Let $p \geqslant 1$, and let l_p be the set of all sequences $\mathbf{x} = (x_n)_{n \geqslant 1}$ in the field $\mathbb{K}$ that are *p-summable,* in the sense that $\sum_{n=1}^{\infty} |x_n|^p$ converges. Equipping l_p with termwise operations of addition and multiplication-by-scalars, prove that

$$\|\mathbf{x}\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}$$

defines a norm on l_p, and that l_p is complete with respect to this l_p-*norm.*

In the case $p = 2$, show that the l_p-norm arises from an inner product, as in Exercise 10.

12. Let S, T be orthogonal linear subspaces of a Hilbert space H such that

$$S + T = \{x + y : x \in S, \ y \in T\}$$

is dense in H. Prove that S and T are both located.

13. Let S be a linear subset of a Hilbert space H such that for each $x \in H$ there exists $y \in S$ with $x - y$ orthogonal to S. Prove that S is closed and located.

14. Construct a Brouwerian counterexample to the statement that every selfadjoint operator on a Hilbert space H is normed. Does every normed operator on H have an adjoint?

15. Construct a Brouwerian counterexample to the statement that every selfadjoint operator on a Hilbert space that has located kernel also has located range.

16. Construct Brouwerian counterexamples to each of the following statements.
    (a) Every bounded linear functional on l_2 (see Exercise 11) is compact.
    (b) If T is a bounded linear mapping of l_1 into $\mathbb{C}^2$ such that $P_i \circ T$ is normed for $i = 1, 2$, then T is compact (where P_i is the ith projection of $\mathbb{C}^2$ onto $\mathbb{C}$).

17. Construct a Brouwerian counterexample to the proposition that the sum of two compact linear mappings between normed spaces is compact.

18. Use the λ-technique as an alternative means of removing the restriction that $\|u\| > 0$ in the proof of the Riesz representation theorem.

19. Let $\mathcal{R}$ be an algebra of normed, jointed operators on a complex Hilbert space H, where the operation of multiplication on $\mathcal{R}$ is just the composition of operators. For each $T \in \mathcal{R}$ and each $\lambda \in \mathbb{C}$ define an operator T' on the direct sum $H \oplus \mathbb{C}$ by

$$T'(x, \zeta) = (Tx + \lambda x, \lambda \zeta) \,.$$

Show that T' is both selfadjoint and normed.

Notes

Our definition of "linearly independent" is classically equivalent to the usual one: namely, if $\sum_{i=1}^{n} \lambda_i e_i = 0$, then $\lambda_i = 0$ for each i. Constructively, the two definitions are equivalent if and only if MP holds. Bishop's definition of "finite-dimensional" requires that the space have a norm from the outset. Our definition is more in the spirit of linear algebra.

Theorem 4.2.1 can be applied to the case of Chebyshev approximation: best approximation of elements of $\mathcal{C}[0,1]$, relative to the sup norm, by polynomials of degree at most n. However, in this special case a deeper analysis enables one to prove the existence of best approximations without the use of Theorem 4.2.1; moreover, that analysis reveals that the best approximation process is, as one would expect in a constructive context, continuous. See [15].

The induction step in the proof of Theorem 4.2.1 is a lot simpler if we are permitted to use Brouwer's fan theorem; for we can then show that any uniformly continuous function from a compact metric space to the positive real numbers has a positive infimum (see [34], Chapter 6).

As Section 4.3 shows, the elements of Hilbert space theory require very few modifications to bring them into constructive line. However, we have to be careful about the requirement that the linear functionals be normed before we can apply the Riesz representation theorem, and about the possibility that a given bounded operator may not have an adjoint.

For a different proof of the Riesz representation theorem see [35]. Working with a direct sum, as was done in the proof of Theorem 4.3.6 in order to circumvent our inability to decide whether a space is trivial, is a useful technique in other applications; see [28].

In the proof that an idempotent, bounded, selfadjoint operator P on a Hilbert space is a projection, we used boundedness only to prove that the set $V = \{x \in H : Px = x\}$ is closed. The Hellinger–Toeplitz theorem (see Chapter 6) shows that every selfadjoint operator P on H is sequentially continuous, a property strong enough to prove that the corresponding set S is closed. Thus, in fact, every idempotent selfadjoint operator on a Hilbert space is a (bounded linear) projection.

5

Linearity and Convexity

Every separation is a link.

—Simone Weil, *'Metaxu', Gravity and Grace*

We begin the chapter by exploring some geometric aspects of convexity that are used later in the construction of one of the cornerstones of functional analysis: the separation theorem. In turn, this leads us to the Hahn–Banach extension theorem, which in its most general form for separable normed spaces produces only approximately norm-preserving extensions of normed linear functionals. We also give Ishihara's version of the Hahn–Banach theorem, which provides a unique norm-preserving extension of a given normed linear functional in the case where the norm on the space is Gâteaux differentiable. We then use the separation and Hahn–Banach theorems to explore the interplay between a normed space and its dual. In particular, we characterise certain linear functionals on spaces of bounded linear mappings. For our discussion of duality we develop the fundamentals of the theory of locally convex topological vector spaces.

5.1 Crossing Boundaries

Suppose we start at a point ξ in the interior of a located subset C of a normed space X and move linearly towards a point z in the metric complement of C. Are we able to tell when we are crossing the boundary

$$\partial C = \overline{C} \cap \overline{\sim C}$$

of C? In general, as is discussed in more detail in problems at the end of this chapter, the constructive answer is no. However, our geometric intuition suggests that when C is convex, we might succeed in pinpointing boundary crossing points. Our first few lemmas are designed to lead us to a proof of the existence of the boundary crossing point, which depends continuously on the origin ξ and the terminus z of our path out of the convex set.

Lemma 5.1.1. *Let C be a convex subset of a normed space X, let $\xi \in C^\circ$, and let $r > 0$ be such that $B(\xi, r) \subset C$. Let $z \neq \xi$, $0 < t < 1$, and $z' = t\xi + (1-t)z$. If the ball $B(z, tr)$ intersects C, then $B(z', t^2 r) \subset C$.*

Proof. Suppose that there exists y in $B(z, tr) \cap C$. Let $\zeta' \in B(z', t^2 r)$ and

$$\xi' = \left(1 - \frac{1}{t}\right) y + \frac{1}{t}\zeta'.$$

Then

$$\begin{aligned}
\|\xi - \xi'\| &= \left\| \left(1 - \frac{1}{t}\right) z + \frac{1}{t} z' - \left(1 - \frac{1}{t}\right) y - \frac{1}{t}\zeta' \right\| \\
&\leqslant \left(\frac{1}{t} - 1\right) \|z - y\| + \frac{1}{t} \|z' - \zeta'\| \\
&< \left(\frac{1}{t} - 1\right) tr + \frac{1}{t} t^2 r \\
&= r.
\end{aligned}$$

Hence $\xi' \in C$. Since $\zeta' = t\xi' + (1-t)y$, it follows by convexity that $\zeta' \in C$. □

In the context of a vector space X over $\mathbb{K}$ we define *intervals* as follows:

$$\begin{aligned}
[x, y] &= \{tx + (1-t)y : 0 \leqslant t \leqslant 1\}, \\
(x, y) &= \{tx + (1-t)y : 0 < t < 1\},
\end{aligned}$$

where $x, y \in X$.

Lemma 5.1.2. *Let C be an open convex subset of a normed space X such that $C \cup -C$ is dense in X, let $\xi \in C$, and let $z \in -C$. Then $(C \cup -C) \cap [\xi, z]$ is dense in $[\xi, z]$.*

Proof. Fix $\varepsilon > 0$ and choose r such that

$$0 < \frac{r}{1+r} \|\xi - z\| < \varepsilon,$$

$B(\xi, r) \subset C$, and $B(z, r) \subset -C$. Given y in $[\xi, z]$, we may assume that $\|y - \xi\| > r/2$ and $\|y - z\| > r/2$, so $y = \alpha\xi + (1-\alpha)z$ for some $\alpha \in (0, 1)$. Fix λ such that

$$\max\left\{\frac{-\varepsilon}{\|\xi - z\|}, \frac{-\alpha}{1-\alpha}\right\} < \lambda < 0, \tag{5.1}$$

and set

$$t = \frac{\lambda}{\lambda - 1},$$
$$y_1 = \lambda\xi + (1-\lambda)\,y.$$

Then $0 < t < 1$ and $y = t\xi + (1-t)\,y_1$. Also,

$$y_1 = [1 - (1-\lambda)(1-\alpha)]\,\xi + (1-\lambda)\,(1-\alpha)\,z,$$

where, by (5.1),

$$0 < (1-\lambda)\,(1-\alpha) < 1,$$

so $y_1 \in [\xi, z]$. Now pick $y' \in C \cup -C$ such that

$$\|y - y'\| < \min\left\{r^2, t^2 r\right\}.$$

Take first the case $y' \in C$. By our choice of r, the point $\xi + \frac{1}{r}(y - y')$ belongs to C. Since C is convex,

$$y'' = \frac{1}{1+r}y' + \frac{r}{1+r}\left(\xi + \frac{1}{r}(y - y')\right)$$

also belongs to C. Moreover,

$$y'' = \frac{r}{1+r}\xi + \frac{1}{1+r}y \in [\xi, y] \subset [\xi, z],$$

and

$$\|y - y''\| = \frac{r}{1+r}\|\xi - y\| \leqslant \frac{r}{1+r}\|\xi - z\| < \varepsilon.$$

We are left with the case $y' \in -C$ to dispose of. But then $B(y, t^2 r)$ intersects $-C$, so, by the preceding lemma, $B(y_1, tr) \cap C = \varnothing$; whence $y_1 \in -C$. Since

$$\|y - y_1\| = |\lambda|\,\|\xi - y\| \leqslant \varepsilon$$

by (5.1), the proof is complete. □

Lemma 5.1.3. *Let X be a normed space, let x_1, x_2 be distinct points of X, and let $x_3 = \lambda x_1 + (1-\lambda)\,x_2$ with $\lambda \neq 0, 1$. For all $\alpha, \beta > 0$, if $\|x - x_1\| < \alpha/|\lambda|$ and $\|y - x_2\| < /\beta/|1-\lambda|$, then $\|\lambda x + (1-\lambda)\,y - x_3\| < \alpha + \beta$.*

Proof. For such x and y we have

$$\|\lambda x + (1-\lambda)\,y - x_3\| \leqslant |\lambda|\,\|x - x_1\| + |1-\lambda|\,\|y - x_2\|,$$

from which the result follows almost immediately. □

In the presence of convexity, the preceding, seemingly innocent, lemma turns out to be a powerful tool. Here is a first example of its use.

Lemma 5.1.4. *If C is an inhabited, open, convex subset of a normed space X, then $-C$ is dense in $\sim C$.*

Proof. Fixing $\xi \in C$, choose $r > 0$ such that $B(\xi, r) \subset C$. Consider any $z \in \sim C$ and $\varepsilon > 0$. Setting

$$t = 1 + \frac{\varepsilon}{\|z - \xi\|}$$

and

$$x_2 = tz + (1-t)\,\xi,$$

we see that $\|x_2 - z\| = \varepsilon$ and that

$$\xi = \frac{-t}{1-t}z + \frac{1}{1-t}x_2.$$

It now suffices to show that $x_2 \in -C$. Taking $x_1 = z, x_3 = \xi$, and $\lambda = t/(t-1)$ in Lemma 5.1.3, we see that if $\|y - x_2\| < r\,(t-1)$, then

$$\left\| \frac{-t}{1-t}z + \frac{1}{1-t}y - \xi \right\| < r$$

and therefore

$$\frac{-t}{1-t}z + \frac{1}{1-t}y \in C.$$

If also $y \in C$, then

$$z = \frac{1}{t}y + \left(1 - \frac{1}{t}\right)\left(\frac{-t}{1-t}z + \frac{1}{1-t}y\right) \in C,$$

a contradiction. Hence $y \notin C$ whenever $\|y - x_2\| < r\,(t-1)$; in other words, $B(x_2, r(t-1)) \subset \neg C$, from which it follows that $x_2 \in -C$. □

A more significant example of the application of Lemma 5.1.3 is the following.

Proposition 5.1.5. *Let C be an open convex subset of a Banach space X such that $C \cup -C$ is dense in X, and let $\xi \in C$. For each $z \in -C$ and each $t \in [0,1]$ write*

$$z_t = t\xi + (1-t)\,z.$$

Then the following hold:

(a) $\gamma(\xi, z) = \inf\{t \in [0,1] : z_t \in C\}$ *exists, and* $0 < \gamma(\xi, z) < 1$.

(b) $z_{\gamma(\xi,z)}$ *is the unique intersection of* $[\xi, z]$ *with* ∂C.

(c) *If* $\gamma(\xi, z) < t \leqslant 1$, *then* $z_t \in C$.

(d) *If* $0 \leqslant t < \gamma(\xi, z)$, *then* $z_t \in -C$.

Moreover, the mapping $(\xi, z) \rightsquigarrow z_{\gamma(\xi,z)}$ *of* $C \times -C$ *into* ∂C *is continuous at each point of* $C \times -C$.

Proof. Fixing z in $-C$, let

$$S = \{t \in [0,1] : z_t \in C\},$$

which contains 1. Note that since C is convex, S is convex also, from which it follows that

(*) if $0 < t' < 1$ and $z_{t'} \notin C$, then $z_t \notin C$ for all t with $0 < t < t'$.

Let $0 \leqslant \alpha < \beta \leqslant 1$. By Lemma 5.1.2, there exists $t_0 \in (\alpha, \beta)$ such that the point $y_0 = t_0\xi + (1 - t_0)\, z$ belongs to $C \cup -C$. Either $t_0 \in S$ or $y_0 \in -C \cap [\xi, z]$. In the latter case, since $\xi \in C \cap [\xi, z]$, we see from (*) that t_0, and hence α, is a lower bound for S. It follows from the least-upper-bound principle that

$$m = \gamma(\xi, z) = \inf S$$

exists. Choosing $r > 0$ such that $\overline{B}(\xi, r) \subset C$ and $\overline{B}(z, r) \subset -C$, now observe that if

$$t = 1 - \frac{r}{\|\xi - z\|},$$

then $0 < t < 1$ and $\|\xi - (t\xi + (1-t)\, z)\| = r$; whence $t\xi + (1-t)\, z \in C$. Thus

$$m \leqslant 1 - \frac{r}{\|\xi - z\|} < 1.$$

On the other hand, since

$$\|z - (1-t)\,\xi + tz\| = r$$

and therefore $z_t \in -C$, we see from (*) that

$$m \geqslant \frac{r}{\|\xi - z\|} > 0.$$

By the definition of m as an infimum, the point

$$h(z) = m\xi + (1-m)\, z$$

belongs to the closure of C. Again applying Lemma 5.1.2, we can find $t' \in (0, m)$ such that $z_{t'}$ belongs to $C \cup -C$ and is arbitrarily close to z_m. By the definition of m, we have $z_{t'} \notin C$, and so $z_{t'} \in -C$. Thus z_m is in the closure of $-C$ and hence in ∂C. The uniqueness part of (b) will follow immediately once we have proved (c) and (d).

Now observe that if $m < t \leqslant 1$, then by the definition of m as an infimum, there exists t' such that $m < t' < t$ and $z_{t'} \in C$; since z_t is in $[\xi, z_{t'}]$, it is in the convex set C. This proves (c). To dispose of (d), let $0 \leqslant t < m$ and let η be the midpoint of the segment $[\xi, z_m]$. Then, by (c), $\eta \in C$, so there exists $\delta > 0$ such that $B(\eta, \delta) \subset C$. Also, $\eta = \lambda z_m + (1-\lambda) z_t$ for some $\lambda > 1$. Consider any point y with $\|y - z_t\| < \delta/(\lambda - 1)$. Applying Lemma 5.1.3 with $x_1 = z_m, x_2 = z_t$, and $x_3 = \eta$, we see that

$$\|\lambda z_m + (1-\lambda) y - \eta\| < \delta$$

and therefore $\lambda z_m + (1-\lambda) y \in C$. If also $y \in C$, then

$$z_m = \frac{1}{\lambda} (\lambda z_m + (1-\lambda) y) + \left(1 - \frac{1}{\lambda}\right) y \in C,$$

which is absurd since $z_m \in \partial C$ and C is open. Thus

$$B\left(z_t, \frac{\delta}{\lambda - 1}\right) \subset \neg C$$

and therefore $z_t \in -C$. This proves (d).

To prove the continuity of the boundary crossing map on $C \times -C$, fix $\xi \in C$, $z \in -C$, and $\varepsilon > 0$. Using (c) and (d), choose a, b, s such that $0 < a < m < b < 1$, $s > 0$,

$$B(z_a, s) \subset -C \cap B(z_m, \varepsilon),$$

and

$$B(z_b, s) \subset C \cap B(z_m, \varepsilon).$$

Consider points $z' \in -C$ and $\xi' \in C$ with

$$\max\{\|z' - z\|, \|\xi' - \xi\|\} < \frac{s}{2} \min\left\{\frac{1}{a}, \frac{1}{b}, \frac{1}{1-a}, \frac{1}{1-b}\right\}.$$

For each $t \in [0,1]$ set

$$z'_t = t\xi' + (1-t) z'.$$

Taking $x_1 = \xi, x_2 = z, x = \xi'$, and $y = z'$, and applying Lemma 5.1.3 with $\lambda = a$, we obtain $\|z'_a - z_a\| < s$ and therefore $z'_a \in -C \cap B(z_m, \varepsilon)$. On the other hand, applying Lemma 5.1.3 with $\lambda = b$, we see that $\|z'_b - z_b\| < s$ and therefore $z'_b \in C \cap B(z_m, \varepsilon)$. It follows that $z'_{\gamma(\xi', x)}$ is in the segment (z'_a, z'_b), which lies in $B(z_m, \varepsilon)$; whence $\|z'_{\gamma(\xi', z')} - z_{\gamma(\xi, z)}\| < \varepsilon$. □

For fixed $\xi \in C$, we call the mapping $z \rightsquigarrow z_{\gamma(\xi, z)}$ in the foregoing proposition the *boundary crossing map of* C *relative to* ξ. We shall use the existence of exact boundary crossings out of convex sets in the next section.

5.2 Separation Theorems

Developing the theme of convexity, in this section we approach the fundamental theorems on the separation of points and convex sets by hyperplanes. In turn, this material will lead us in the next section to the Hahn–Banach theorem, one of the cornerstones of functional analysis.

A subset C of a vector space X over $\mathbb{K}$ is called a *cone* if for all $x, y \in C$ and all $t > 0$, both $x + y$ and tx belong to C. In that case, C is convex. The closure of a cone is a cone, as is the intersection of two cones.

If K is a convex subset of X, then the set

$$c(K) = \{tx : x \in K, t > 0\}$$

is a cone. For clearly, if $x \in c(K)$, then $tx \in c(K)$ for all $t > 0$; whereas if $x_1, x_2 \in c(K)$, then $x_1 = t_1 y_1$ and $x_2 = t_2 y_2$ for some $y_1, y_2 \in K$ and some $t_1, t_2 > 0$, so $x_1 + x_2 = (t_1 + t_2) z$ with

$$z = \frac{t_1}{t_1 + t_2} y_1 + \frac{t_2}{t_1 + t_2} y_2 \in K.$$

We call $c(K)$ the *cone generated by the convex set* K. If X is a normed space and K is open, then so is $c(K)$.

Lemma 5.2.1. *Let K be a bounded located subset of a normed space X, $a \in X$, and let $\tau > 0$. Then the set*

$$S = \{tx + (1 - t) a : x \in K, 0 < t < \tau\}$$

is located.

Proof. Choose $R > 0$ such that $\|x\| \leqslant R$ for all $x \in K$. Fixing $x_0 \in X$, note first that for each $t > 0$,

$$f(t) = \rho(x_0, \{tx + (1 - t) a : x \in K\})$$

exists and equals

$$t\rho\left(\frac{1}{t} x_0 + \frac{t - 1}{t} a, K\right).$$

If also $t' > 0$, then since

$$\|x_0 - tx - (1 - t) a\| \leqslant \|x_0 - t'x - (1 - t') a\| + (\|a\| + \|x\|) |t - t'|$$

for all $x \in X$, we see that

$$f(t) \leqslant f(t') + (\|a\| + R) |t - t'|.$$

It follows that the mapping f is uniformly continuous on $\mathbb{R}^+$; whence $\rho(x_0, S)$ exists as the infimum of f on the totally bounded interval $(0, \tau)$. $\square$

Lemma 5.2.2. *Let K be a bounded, located, convex subset of a normed space X such that $\rho(0,K) > 0$, and let $a \in X$ be such that $\rho(-a, c(K)) > 0$. Then there exists $r > 0$ such that $\|tx + (1-t)\,a\| \geqslant r$ for all $x \in K$ and $t > 0$.*

Proof. Let
$$\delta = \frac{\rho(0,K)}{3\left(\|a\| + \rho(0,K)\right)}.$$
Given x in K and $t > 0$, we have either $|1-t| > \delta$ or $|1-t| < 2\delta$. In the first case,
$$\|tx + (1-t)\,a\| = |1-t| \left\| \frac{t}{1-t}x + (-a) \right\| > \delta\rho(-a, c(K)).$$
In the case $|1-t| < 2\delta$,
$$\begin{aligned} \|tx + (1-t)\,a\| &\geqslant t\,\|x\| - |1-t|\,\|a\| \\ &\geqslant (1-2\delta)\,\rho\,(0,K) - 2\delta\,\|a\| \\ &= \frac{1}{3}\rho(0,K). \end{aligned}$$
Setting
$$r = \min\left\{ \delta\rho(-a, c(K)), \frac{1}{3}\rho(0,K) \right\} > 0,$$
we see that $\|tx + (1-t)\,a\| \geqslant r$ in either case. □

Lemma 5.2.3. *Let K be a bounded, located, convex subset of a normed space X such that $\rho(0,K) > 0$. Then $c\,(K)$ is located.*

Proof. Fix $x_0 \in X$. For each $t > 0$ and each $x \in K$ we have
$$\|x_0 - tx\| \geqslant |t|\,\|x\| - \|x_0\|\,.$$
Hence
$$\rho\,(x_0, tK) \geqslant t\rho\,(0,K) - \|x_0\| \longrightarrow \infty \text{ as } t \longrightarrow \infty.$$
Compute $\tau > 1$ such that
$$\rho\,(x_0, tK) > \rho\,(x_0, K) \qquad (t > \tau - 1)\,. \tag{5.2}$$
Then
$$d = \inf\left\{\rho\,(x_0, tK) : 0 < t < \tau\right\}$$
exists, by the case $a = 0$ of Lemma 5.2.1. Since $\tau > 1$ and therefore $d \leqslant \rho\,(x_0, K)$, it follows from (5.2) that $\rho\,(x_0, c\,(K))$ exists and equals d. □

Lemma 5.2.4. *Let K and L be open cones in a normed space X such that $K \cup L$ is dense in X and $K \subset {\sim}L$. Then*

(a) *$K \subset -L$ and $L \subset -K$,*

(b) *$K \cup -K$ and $L \cup -L$ are dense in X, and*

(c) *K and L have a common boundary—namely, $\overline{K} \cap \overline{L}$.*

If also $L = \{-x : x \in K\}$, then ∂K is a subspace of X.

Proof. Since $L \subset {\sim}{\sim}L \subset {\sim}K$, we see that $K \cup {\sim}K$ is dense; whence, by Lemma 5.1.4, $K \cup -K$ is dense. On the other hand, $K \subset {\sim}L$ and K is open, so $K \subset -L$ and therefore $\overline{K} \subset \overline{-L}$. But $-L$ is open and $K \cup L$ is dense, so $-L \subset \overline{K}$ and therefore, by Lemma 5.1.4, $\overline{{\sim}L} = \overline{-L} \subset \overline{K}$. Hence

$$\overline{K} = \overline{-L} = \overline{{\sim}L}.$$

Interchanging the roles of K and L, we obtain $L \cup -L$ dense and

$$\overline{L} = \overline{-K} = \overline{{\sim}K}.$$

We now have

$$\partial K = \overline{K} \cap \overline{{\sim}K} = \overline{K} \cap \overline{L} = \overline{{\sim}L} \cap \overline{L} = \partial L.$$

Since K, L, and therefore their closures are all cones, so is ∂K.

Now suppose also that $L = \{-x : x \in K\}$, and consider $x \in \partial K = \overline{K} \cap \overline{L}$. There exist points $y \in K$ and $z \in L$ arbitrarily close to x. Then $-y \in L$, $-z \in K$, and these two points are arbitrarily close to $-x$; whence $-x$ belongs to ∂K. It follows that ∂K is a subspace of X. □

Lemma 5.2.5. *Let K be a located subset of a normed space X, and let $r > 0$. Then the set*

$$K_r = \{x \in X : \rho(x, K) \leqslant r\}$$

is located, and for each $x_0 \in X$,

$$\rho(x_0, K_r) = \max\{0, \rho(x_0, K) - r\}. \tag{5.3}$$

Proof. If $x \in K_r$ and $\varepsilon > 0$, then there exists $y \in K$ such that $\|x - y\| < r + \varepsilon$; so

$$\|x_0 - x\| \geqslant \|x_0 - y\| - \|x - y\| \geqslant \rho(x_0, K) - r - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we have

$$\|x_0 - x\| \geqslant \rho(x_0, K) - r \qquad (x \in K_r). \tag{5.4}$$

On the other hand, choosing $z \in K$ such that $\|x_0 - z\| < \rho(x_0, K) + \varepsilon$, we have either $\|x_0 - z\| < r$, in which case (5.3) holds with each side equal to 0, or else $\|x_0 - z\| > 0$. In the latter case, writing

$$\zeta = \min\left\{1, \frac{r}{\|x_0 - z\|}\right\}(x_0 - z),$$

we have $z + \zeta \in K_r$ and

$$\begin{aligned}\|x_0 - (z+\zeta)\| &= \left\|\left(1 - \min\left\{1, \frac{r}{\|x_0 - z\|}\right\}\right)(x_0 - z)\right\| \\ &= \max\{0, \|x_0 - z\| - r\} \\ &< \max\{0, \rho(x_0, K) - r\} + \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, it follows from this and (5.4) that (5.3) holds. □

By a *half-space* of a normed space X we mean a convex subset K such that ∂K is a hyperplane and the set

$$\{x \in X : x \in K \vee -x \in K\}$$

is dense in X.

We are now ready for the *basic separation theorem*. Our proof illustrates an important observation about classical proofs using Zorn's lemma (see [65]): for separable spaces it is often possible to replace such a proof by a constructive one that uses an induction argument.

Theorem 5.2.6. *Let X be a separable normed space, K_0 a bounded, located, open, convex subset of X such that $\rho(0, K_0) > 0$, and x_0 a point of X such that $-x_0 \in K_0$. Then there exists an open half-space K of X such that $K_0 \subset K$, $\rho(x_0, K) > 0$, and the boundary of K is a located subspace of X that is a hyperplane with associated vector x_0.*

Proof. Let $(x_n)_{n\geqslant 1}$ be a dense sequence in X. The basic idea of the proof is this: we carry out a succession of located convex enlargements of K_0 such that for $n \geqslant 1$, the cone generated by the nth enlargement K_n is close to at least one of the points x_n and $-x_n$, and such that the union of the cones $c(K_n)$ is the desired open half-space. The idea may seem simple, but the details are rather complicated.

To be more precise, we construct bounded, located, open, convex subsets $K_0 \subset K_1 \subset K_2 \subset \cdots$ of X such that the following properties obtain for $n \geqslant 1$:

(a) $\rho(x_0, c(K_n)) > (1 - 2^{-n})\,\rho(x_0, c(K_{n-1}))$,

(b) $\rho(0, K_n) > 0$,

(c) $\max\left\{\rho\left(x_n, c(K_n)\right), \rho\left(-x_n, c(K_n)\right)\right\} < \dfrac{1}{n}$.

To that end, assume that $K_0, \ldots, K_{n-1}$ have been constructed with the applicable properties. Then $c(K_{n-1})$ is located, by Lemma 5.2.3. Either

$$\min\left\{\rho(x_n, c(K_{n-1})), \rho(-x_n, c(K_{n-1}))\right\} < \frac{1}{n},$$

in which case we take $K_n = K_{n-1}$, or else both the numbers $\rho(x_n, c(K_{n-1}))$ and $\rho(-x_n, c(K_{n-1}))$ are greater than $1/2n$. In the latter case, writing

$$\begin{aligned} K_n^+ &= \left\{tx + (1-t)\,x_n : 0 < t < 1,\ x \in K_{n-1}\right\}, \\ K_n^- &= \left\{tx - (1-t)\,x_n : 0 < t < 1,\ x \in K_{n-1}\right\}, \end{aligned}$$

we obtain two bounded, open, convex sets that, by Lemma 5.2.1, are located. We show that at least one of the sets K_n^+, K_n^- satisfies our requirements for K_n. Since K_{n-1} is open, for each $x \in K_{n-1}$ we can find $\lambda > 1$ such that $\lambda x - (\lambda - 1)\,x_n \in K_{n-1}$; so, by convexity,

$$x = \left(1 - \frac{1}{\lambda}\right) x_n + \frac{1}{\lambda}\left(\lambda x - (\lambda - 1)\,x_n\right) \in K_n^+.$$

Hence $K_{n-1} \subset K_n^+$; likewise, $K_{n-1} \subset K_n^-$. Since, by our hypotheses, $\rho(0, K_{n-1}) > 0$ and $\rho(-x_n, c\left(K_{n-1}\right)) > 1/2n$, we see from Lemma 5.2.2 that $\rho(0, K_n^+) > 0$; similarly, $\rho\left(0, K_n^-\right) > 0$.

Next, given $z^+ \in c(K_n^+)$ and $z^- \in c(K_n^-)$, find $t^+, t^- > 0$ and $x^+, x^- \in c(K_{n-1})$ such that $z^+ = t^+ x_n + x^+$ and $z^- = -t^- x_n + x^-$. We have

$$\begin{aligned} t^-&\left\|x_0 - z^+\right\| + t^+\left\|x_0 - z^-\right\| \\ &= t^-\left\|x_0 - \left(t^+ x_n + x^+\right)\right\| + t^+\left\|x_0 - \left(-t^- x_n + x^-\right)\right\| \\ &\geqslant \left\|\left(t^+ + t^-\right) x_0 - \left(t^- x^+ + t^+ x^-\right)\right\| \\ &= \left(t^+ + t^-\right)\left\|x_0 - \left(\frac{t^-}{t^+ + t^-} x^+ + \frac{t^+}{t^+ + t^-} x^-\right)\right\| \\ &\geqslant \left(t^+ + t^-\right)\rho\left(x_0, c(K_{n-1})\right) \\ &> \left(t^+ + t^-\right)\left(1 - 2^{-n-1}\right)\rho\left(x_0, c(K_{n-1})\right). \end{aligned}$$

Hence either

$$t^-\left(\left\|x_0 - z^+\right\| - \left(1 - 2^{-n-1}\right)\rho\left(x_0, c(K_{n-1})\right)\right) > 0$$

or

$$t^+\left(\left\|x_0 - z^-\right\| - \left(1 - 2^{-n-1}\right)\rho\left(x_0, c(K_{n-1})\right)\right) > 0,$$

from which we obtain

$$\max\left\{\left\|x_0 - z^+\right\|, \left\|x_0 - z^-\right\|\right\} > \left(1 - 2^{-n-1}\right)\rho\left(x_0, c(K_{n-1})\right).$$

Taking the infimum as z^+, z^- run over $c(K_n^+), c(K_n^-)$ respectively, we obtain

$$\begin{aligned}\max\left\{\rho(x_0, c(K_n^+)), \rho(x_0, c(K_n^-))\right\} &\geqslant \left(1-2^{-n-1}\right)\rho\left(x_0, c(K_{n-1})\right)\\ &> \left(1-2^{-n}\right)\rho\left(x_0, c(K_{n-1})\right).\end{aligned}$$

Thus either

$$\rho(x_0, c(K_n^+)) > \left(1-2^{-n}\right)\rho\left(x_0, c(K_{n-1})\right),$$

in which case, noting that $x_n \in \overline{K_n^+}$ and therefore $\rho\left(x_n, c(K_n^+)\right) = 0 < 1/n$, we take $K_n = K_n^+$; or else

$$\rho(x_0, c(K_n^-)) > \left(1-2^{-n}\right)\rho\left(x_0, c(K_{n-1})\right),$$

when we take $K_n = K_n^-$. This completes the inductive construction of K_n.

The set

$$K = \bigcup_{n=0}^{\infty} c(K_n)$$

is easily shown to be an open convex cone. Moreover, we see from (a) that $(\rho(x_0, c(K_n)))_{n\geqslant 1}$ is a Cauchy sequence, which therefore converges in $\mathbb{R}$. Since we have $c(K_{n-1}) \subset c(K_n)$ for each n, the limit of the sequence is $\rho(x_0, K)$. Now, for $n \geqslant 1$,

$$\prod_{k=1}^{n}\left(1-2^{-k}\right) \geqslant \frac{1}{2}\left(1-\sum_{k=2}^{n} 2^{-k}\right) > \frac{1}{4},$$

so, by (a),

$$\rho(x_0, K) = \lim_{n\to\infty}\rho\left(x_0, c\left(K_n\right)\right) \geqslant \frac{1}{4}\rho\left(x_0, c\left(K_0\right)\right).$$

But by the convexity of K_0,

$$\begin{aligned}\rho\left(x_0, c(K_0)\right) &= \inf\left\{\|tx - x_0\| : t > 0,\ x \in K_0\right\}\\ &= \inf\left\{(t+1)\left\|\frac{t}{t+1}x + \frac{1}{t+1}\left(-x_0\right)\right\| : t > 0,\ x \in K_0\right\}\\ &\geqslant \rho(0, K_0).\end{aligned}$$

Hence $\rho(x_0, K) > 0$.

Let

$$L = \{x \in X : -x \in K\}.$$

Then L is an open convex cone containing x_0. By (c), for each n, either $\rho\left(x_n, K\right) < 1/n$ or $\rho\left(x_n, L\right) < 1/n$. So $K \cup L$ is dense in X. On the other hand, $K \subset {\sim}L$: for if $x \in K$ and $y \in L$, then there exists n such that $x \in c\left(K_n\right)$ and $-y \in c\left(K_n\right)$; so there exist $t > 0$ and $z \in K_n$ such that $x - y = tz$ and therefore $\|x-y\| \geqslant t\rho\left(0, K_n\right) > 0$. It now follows from Lemma 5.2.4 that $L \cup -L$ is dense in X, and the common boundary of K and L is the set

$$N = \overline{K} \cap \overline{L},$$

which is a linear subspace of X.

To prove that N is a hyperplane, let γ denote the boundary crossing map of L relative to its interior point x_0. Note that, by Lemma 5.2.4, $K \subset -L$; so, by Proposition 5.1.5, for each $x \in K$ there exists a unique $t \in (0,1)$ such that

$$\gamma(x) = tx_0 + (1-t)\,x$$

and therefore

$$x = \frac{t}{1-t}(-x_0) + \frac{1}{1-t}\gamma(x).$$

Similar considerations using the boundary crossing map of K relative to $-x_0$ complete a proof that each $x \in K \cup L$ can be expressed in the form $\alpha x_0 + \beta z$ with $\alpha \in \mathbb{R}$ and $z \in N$. In fact, such an expression obtains for all $x \in X$. For since $K \cup L$ is dense, there exist a sequence $(\alpha_n)_{n \geqslant 1}$ in $\mathbb{R}$ and a sequence $(z_n)_{n \geqslant 1}$ in N such that $\alpha_n x_0 + z_n \longrightarrow x$ as $n \longrightarrow \infty$. Given $\varepsilon > 0$ and $m > n$, we have either $|\alpha_m - \alpha_n| < \varepsilon$ or $|\alpha_m - \alpha_n| > 0$; in the latter case, since $z_m - z_n \in N \subset \overline{K}$,

$$\begin{aligned}\|(\alpha_m x_0 + z_m) - (\alpha_n x_0 + z_n)\| &= |\alpha_m - \alpha_n| \left\| x_0 - \frac{1}{|\alpha_m - \alpha_n|}(z_m - z_n) \right\| \\ &\geqslant |\alpha_m - \alpha_n|\,\rho(x_0, K),\end{aligned}$$

so

$$|\alpha_m - \alpha_n| \leqslant \frac{1}{\rho(x_0,K)} \|(\alpha_m x_0 + z_m) - (\alpha_n x_0 + z_n)\|.$$

It follows that $(\alpha_n)_{n \geqslant 1}$ is a Cauchy sequence and therefore converges to a limit α in $\mathbb{R}$. Writing $z = x - \alpha x_0$, we have

$$\begin{aligned}\|z - z_n\| &\leqslant \|x - (\alpha_n x_0 + z_n)\| + |\alpha - \alpha_n|\,\|x_0\| \\ &\longrightarrow 0 \text{ as } n \longrightarrow \infty,\end{aligned}$$

so $z \in \overline{N} = N$. Thus N is a hyperplane with associated vector x_0, and therefore (since also $K \cup L$ is dense in X) K is a half-space.

Now, for each $x \in K$ we have $\gamma(x) \in N \subset \overline{K}$ and $\|x_0 - \gamma(x)\| \leqslant \|x_0 - x\|$. It follows that $\rho(x_0, N)$ exists, equals $\rho(x_0, K)$, and is positive. For each $x \in X$, choosing $\alpha \in \mathbb{R}$ and $z \in N$ such that $x = \alpha x_0 + z$, we easily see that $\rho(x, N)$ exists and equals $|\alpha|\,\rho(x_0, K)$. Hence N is located. □

Corollary 5.2.7. *Under the hypotheses of* Theorem 5.2.6, *there exists a normed linear functional u on X such that $\|u\| = 1$ and* $\mathrm{Re}\,u$ *is positive on K_0.*

Proof. First take the case $\mathbb{K} = \mathbb{R}$. Let K be as in the conclusion of Theorem 5.2.6, let

$$L = \{-x : x \in K\},$$

recall that $K \subset -L$, and let N be the common boundary of K and L. Since N is a located hyperplane, we see from Propositions 2.3.4 and 2.3.6 that there exists a normed linear functional v on X such that $\ker(v) = N$ and $v(x_0) = 1$. Defining $u = -\|v\|^{-1} v$, we obtain a normed linear functional on X with norm 1 such that $\ker(u) = N$ and $u(x_0) < 0$. Let γ be the boundary crossing map of L relative to x_0. Given $x \in K$, choose $\lambda \in (0,1)$ such that $\gamma(x) = \lambda x + (1-\lambda) x_0$. Then $u(\gamma(x)) = 0$, so

$$u(x) = -\frac{1-\lambda}{\lambda} u(x_0) > 0.$$

Hence u is positive on K and therefore on K_0.

Now consider the case $\mathbb{K} = \mathbb{C}$. By the case just considered, there exists a real linear functional $v : X \longrightarrow \mathbb{R}$ on the real linear space X such that v has norm 1 and is positive on K_0. The following lemma then shows that

$$u(x) = v(x) - \mathrm{i}v(\mathrm{i}x)$$

defines a linear functional $u : X \longrightarrow \mathbb{C}$ on the complex linear space X such that $\|u\| = 1$ and $\operatorname{Re} u$ is positive on K_0. □

Lemma 5.2.8. *If X is a normed linear space over $\mathbb{C}$, then there is a one-one correspondence between real linear functionals v on X and complex linear functionals u on X, given by*

$$u(x) = v(x) - \mathrm{i}v(\mathrm{i}x). \tag{5.5}$$

Moreover, if either u or v is normed, then both are and their norms are equal.

Proof. If u is a complex linear functional, then $u = v + \mathrm{i}w$ for unique real linear functionals v and w. But

$$\mathrm{i}v(x) - w(x) = \mathrm{i}u(x) = u(\mathrm{i}x) = v(\mathrm{i}x) + \mathrm{i}w(\mathrm{i}x),$$

so $w(x) = -v(\mathrm{i}x)$.

Conversely, if v is a real linear functional, then (5.5) defines a linear functional u that respects multiplication by real numbers; it is easily checked that $u(\mathrm{i}x) = \mathrm{i}u(x)$. The final conclusion, about the norms of u and v, is a straightforward consequence of the identity

$$|u(x)| = \sup\{|v(\lambda x)| : |\lambda| = 1\}, \tag{5.6}$$

which we now establish. If $|\lambda| = 1$, then

$$|u(x)| = |\lambda u(x)| = |u(\lambda x)| \geqslant |\operatorname{Re} u(\lambda x)| = |v(\lambda x)|.$$

On the other hand, if $\varepsilon > 0$, then either $|u(x)| - \varepsilon < |v(x)|$, or else $u(x) \neq 0$. In the latter case, if we let $\lambda = |u(x)|/u(x)$, then

$$u(\lambda x) = \lambda u(x) = |u(x)| \in \mathbb{R},$$

so $|u(x)| = \operatorname{Re} u(\lambda x) = v(\lambda x)$. Notice that in this case the supremum in (5.6) is achieved. □

The following result is the full form of the *separation theorem.*

Theorem 5.2.9. *Let A and B be bounded convex subsets of a separable normed space X such that the* **algebraic difference**

$$\{y - x : x \in A,\ y \in B\}$$

is located and the **mutual distance**

$$d = \inf \{\|y - x\| : x \in A,\ y \in B\}$$

is positive. Then for each $\varepsilon > 0$ there exists a normed linear functional u on X, with norm 1, such that

$$\operatorname{Re} u(y) > \operatorname{Re} u(x) + d - \varepsilon \quad (x \in A, y \in B)\,.$$

Proof. We may assume that $\varepsilon < d$. Write

$$K_0 = \left\{y - x - z : x \in A,\ y \in B,\ \|z\| < d - \frac{\varepsilon}{2}\right\},$$

which is bounded, open, and convex. It follows from Lemma 5.2.5 that K_0 is located and $\rho(0, K_0) = \varepsilon/2$. Hence, by Corollary 5.2.7, there exists a normed linear functional u on X such that $\|u\| = 1$ and $\operatorname{Re} u$ is positive on K_0. Choose $z \in X$ such that $\|z\| < d - \varepsilon/2$ and $u(z) > d - \varepsilon$; this is possible since $\|u\| = 1$. For all $x \in A$ and $y \in B$ we have $y - x - z \in K_0$ and so $\operatorname{Re} u(y - x - z) > 0$; whence

$$\operatorname{Re} u(y) > \operatorname{Re}\,(u(x) + u(z)) > \operatorname{Re} u(x) + d - \varepsilon,$$

as required. □

An inhabited subset K of a normed space X is said to be *balanced* if $\alpha x \in K$ whenever $x \in K$ and $|\alpha| \leqslant 1$; in that case, $0 \in K$.

Corollary 5.2.10. *Let x_0 be a vector in a separable normed space X, and K a located subset of X that is bounded, convex, and balanced, such that $\rho(x_0, K) > 0$. Then for each $\varepsilon > 0$ there exists a normed linear functional u on X with norm 1 such that*

$$u(x_0) > |u(x)| + \rho(x_0, K) - \varepsilon$$

for all $x \in K$.

Proof. We may assume that $0 < \varepsilon < \rho(x_0, K)$. Applying Theorem 5.2.9 with $A = K$ and $B = \{x_0\}$, construct a normed linear functional v on X with norm 1 such that

$$\operatorname{Re} v(x_0) > \operatorname{Re} v(x) + \rho(x_0, K) - \varepsilon \quad (x \in K).$$

Since $0 \in K$, we have $v(x_0) \neq 0$. Thus

$$u = \frac{|v(x_0)|}{v(x_0)} v$$

is a normed linear functional on X with norm 1 such that

$$u(x_0) = |v(x_0)| \geqslant \operatorname{Re} v(x_0) > \operatorname{Re} v(x) + \rho(x_0, K) - \varepsilon \quad (x \in K).$$

For each $x \in K$, either $u(x_0) > |u(x)| + \rho(x_0, K) - \varepsilon$, or else $u(x) \neq 0$ and therefore $v(x) \neq 0$. In the latter case, since K is balanced,

$$\frac{|v(x)|}{v(x)} x \in K,$$

and therefore

$$\begin{aligned} u(x_0) &> \operatorname{Re} v\left(\frac{|v(x)|}{v(x)} x\right) + \rho(x_0, K) - \varepsilon \\ &= |v(x)| + \rho(x_0, K) - \varepsilon \\ &= |u(x)| + \rho(x_0, K) - \varepsilon, \end{aligned}$$

as required. □

5.3 The Hahn–Banach Theorem

The Hahn–Banach theorem enables us to extend a normed linear functional, with an arbitrarily small increase in norm, from a subspace of a normed space to the entire space. This fundamental result has numerous applications throughout functional analysis.

In the constructive context we deal only with the extension of linear functionals on subspaces of a separable normed space. The standard classical proofs extending the theorem to nonseparable normed spaces depend on Zorn's lemma and are therefore nonconstructive.

We begin with two more consequences of the separation theorem.

Proposition 5.3.1. *Let x be an element of a nontrivial separable normed space X, and let $\varepsilon > 0$. Then there exists a normed linear functional u on X such that $\|u\| = 1$ and $u(x) > \|x\| - \varepsilon$.*

Proof. If $x \neq 0$, then we may apply the separation theorem (Theorem 5.2.9) with $A = \{0\}$ and $B = \{x\}$. In the general case, choose a nonzero vector y such that $\|x - y\| < \varepsilon/2$, and then construct a normed linear functional u on X such that $\|u\| = 1$ and $u(y) > \|y\| - \varepsilon/2$. We have

$$u(x) \geqslant u(y) - |u(x) - u(y)| > \|y\| - \frac{\varepsilon}{2} - \|x - y\| > \|x\| - \varepsilon,$$

as required. □

Corollary 5.3.2. *Let V be a located subspace of a separable normed space X such that the metric complement $-V$ is inhabited. Then there exists a normed linear functional u on X such that $\|u\| = 1$ and $u(V) = \{0\}$.*

Proof. Let e be a unit vector in $-V$. Applying Proposition 5.3.1 in the quotient space X/V, construct a normed linear functional u on X/V such that $\|u\| = 1$ and $u(v) \neq 0$. Clearly, $V \subset \ker(u)$. Since $\| \ \|_{X/V} \leqslant \| \ \|$, we see that u, regarded as a linear functional on the original normed space X, has bound 1. On the other hand, given $\varepsilon > 0$, we can find $x \in X$ with $\|x\|_{X/V} < 1$ and $u(x) > 1 - \varepsilon$; choosing $y \in V$ with $\|x - y\| < 1$, we then have $u(x - y) = u(x) > 1 - \varepsilon$. Since $\varepsilon > 0$ is arbitrary, it follows that, as a linear functional on X, u is normed and has norm 1. □

Proposition 5.3.1 is crucial for our proof of the *Hahn–Banach theorem.*

Theorem 5.3.3. *Let v be a nonzero bounded linear functional on a linear subset Y of a separable normed linear space X such that $\ker(v)$ is located in X. Then for each $\varepsilon > 0$ there exists a normed linear functional u on X such that $\|u\| \leqslant \|v\| + \varepsilon$ and $u(y) = v(y)$ for each $y \in Y$.*

Proof. First note that, by Proposition 2.3.6, v is normed as a linear functional on Y. Fix y_0 in Y with $v(y_0) = 1$. Then for each $x \in \ker(v)$,

$$\|y_0 - x\| \geqslant \frac{1}{\|v\|} v(y_0 - x) = \frac{1}{\|v\|},$$

from which it follows that

$$\|y_0\|_{X/\ker(v)} = \rho\left(y_0, \ker(v)\right) \geqslant \frac{1}{\|v\|}.$$

For each normed linear functional u on $X/\ker(v)$ denote the norm by $\|u\|_{X/\ker(v)}$. Using Proposition 5.3.1, construct a normed linear functional u_0 on $X/\ker(v)$ such that $\|u_0\|_{X/\ker(v)} = 1$ and $u_0(y_0) > 1/(\|v\| + \varepsilon)$. Then

$$u = \frac{1}{u_0(y_0)} u_0$$

is a normed linear functional on $X/\ker(v)$ such that $\|u\|_{X/\ker(v)} < \|v\| + \varepsilon$ and $u(y_0) = 1$. Since $\|\ \|_{X/\ker(v)} \leqslant \|\ \|$ on X, we see that u, regarded as a linear functional on the original normed space X, has bound $\|u\|_{X/\ker(v)}$. To see that this is actually the norm of u on X, consider any $\delta > 0$ and choose $x \in X$ such that $\|x\|_{X/\ker(v)} < 1$ and $u(x) > \|u\|_{X/\ker(v)} - \delta$. Finding $y \in \ker(v)$ such that $\|x - y\| < 1$, we have

$$u(x - y) = u(x) > \|u\|_{X/\ker(v)} - \delta.$$

Since $\delta > 0$ is arbitrary, we conclude that u is normed on X, with

$$\|u\| = \|u\|_{X/\ker(v)} \leqslant \|v\| + \varepsilon.$$

Finally, for each $y \in Y$ we have $y - v(y)y_0 \in \ker(v)$, so

$$0 = u(y - v(y)y_0) = u(y) - v(y),$$

and therefore $u(y) = v(y)$. □

The Hahn–Banach theorem and some of its associates have surprising applications. As surprising as any is the following, whose classical proof is almost trivial.

Proposition 5.3.4. *Let $x_1, \ldots, x_n$ be elements of an infinite-dimensional normed space X, and let $\varepsilon > 0$. Then there exist linearly independent elements $e_1, \ldots, e_n$ of X such that $\|x_i - e_i\| < \varepsilon$ for each i.*

Proof. Construct a finite-dimensional subspace V of span$\{x_1, \ldots, x_n\}$ such that for each i there exists $y_i \in V$ with $\|x_i - y_i\| < \varepsilon/2$ (see Lemma 4.1.11). Since X is infinite-dimensional and V is finite-dimensional, we can embed V in an n-dimensional subspace W of X. We may assume that $y_1 \neq 0$. Setting $e_1 = y_1$, suppose that for some $k < n$, we have constructed linearly independent elements $e_1, \ldots, e_k$ of W such that $\|y_i - e_i\| < \varepsilon/2$ for $1 \leqslant i \leqslant k$. Let V_k be the k-dimensional subspace of W with basis $\{e_1, \ldots, e_k\}$. By Corollary 5.3.2, there exists a normed linear functional u on W such that $u(V_k) = \{0\}$ and $\|u\| = 1$. Construct a vector $z \in W$ such that $\|z\| = \varepsilon/2$ and $u(z) > \varepsilon/3$. Either $u(y_{k+1}) \neq 0$ or else $|u(y_{k+1})| < \varepsilon/3$. In the first case, $\rho(y_{k+1}, V_k) > 0$ and we set $e_{k+1} = y_{k+1}$. In the second case, $u(y_{k+1} - z) \neq 0$, $\rho(y_{k+1} - z, V_k) > 0$, and we set $e_{k+1} = y_{k+1} - z$. In each case, the vectors $e_1, \ldots, e_{k+1}$ are linearly independent, by Lemma 4.1.10, and $\|x_{k+1} - e_{k+1}\| < \varepsilon$. This completes the inductive construction. □

The classical Hahn–Banach theorem says that we can extend a bounded linear functional v from Y to a functional u on the whole space X with exact preservation of norm: that is, $\|u\| = \|v\|$. In general, as Exercise 5 shows, we cannot do this constructively. But, as we shall see, if we impose extra conditions on the norm of X, then we can make the extension norm-preserving.

A mapping u of a linear space X into $\mathbb{R}$ is said to be *convex* if

$$u\left(\lambda x+(1-\lambda)\,y\right)\leqslant \lambda u(x)+(1-\lambda)\,u(y)$$

whenever $x, y \in X$ and $0 \leqslant \lambda \leqslant 1$.

Lemma 5.3.5. *Let u be a convex mapping of a linear space X into $\mathbb{R}$ such that $u(x) = -u(-x)$ for each $x \in X$. Then u is linear.*

Proof. First note that $u(0) = -u(-0) = -u(0)$, so $u(0) = 0$. For each $x \in X$ and for $0 \leqslant \lambda \leqslant 1$ we have

$$u(\lambda x) = u(\lambda x + (1-\lambda)\,0) \leqslant \lambda u(x) + (1-\lambda)\,u(0) = \lambda u(x).$$

Replacing x by $-x$, we obtain $u(-\lambda x) \leqslant \lambda u(-x)$; whence, by our hypotheses on u, $-u(\lambda x) \leqslant -\lambda u(x)$ and therefore $\lambda u(x) \leqslant u(\lambda x)$. Hence $u(\lambda x) = \lambda u(x)$ whenever $0 \leqslant \lambda \leqslant 1$.

Now consider any real $\lambda \neq 0$. If $\lambda > 0$, then by the foregoing,

$$\frac{\lambda}{\lambda+1}u(x) = u\left(\frac{\lambda}{\lambda+1}x\right) = \frac{1}{\lambda+1}u(\lambda x)$$

and therefore $\lambda u(x) = u(\lambda x)$. If $\lambda < 0$, then

$$u(\lambda x) = u((-\lambda)(-x)) = -\lambda u(-x) = \lambda(-u(-x)) = \lambda u(x).$$

Now consider any $\lambda \in \mathbb{R}$, and assume that $u(\lambda x) \neq \lambda u(x)$. The foregoing shows that

$$\neg\,(\lambda > 0 \vee \lambda < 0)$$

and therefore $\lambda = 0$. But this is absurd, since $u(0x) = 0 = 0u(x)$. We conclude that $u(\lambda x) = \lambda u(x)$ for all $\lambda \in \mathbb{R}$.

It remains to prove the additivity of u. By the convexity of u,

$$u\left(\frac{1}{2}(x+y)\right) \leqslant \frac{1}{2}u(x) + \frac{1}{2}u(y).$$

Applying the first part of the proof yields $u(x+y) \leqslant u(x) + u(y)$. We now replace x, y by $-x, -y$ respectively, to obtain

$$u(-(x+y)) = u(-x-y) \leqslant u(-x) + u(-y)$$

and therefore, by the hypotheses on u, $u(x+y) \geqslant u(x) + u(y)$. Hence $u(x+y) = u(x) + u(y)$. □

The norm of a normed space X is said to be *Gâteaux differentiable at* $x \in X$ if for each $y \in X$ the limit

$$u_x(y) = \lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t}$$

exists. For each $c > 0$ we then have

$$u_{cx}(y) = \lim_{t \to 0} \frac{\left\|x + c^{-1}ty\right\| - \|x\|}{c^{-1}t} = u_x(y).$$

If the norm of X is Gâteaux differentiable at each nonzero point of X, we say that X has *Gâteaux differentiable norm.* In such a case we obtain a strengthening of Proposition 5.3.1.

Proposition 5.3.6. *Let x be a nonzero vector in a normed space X, and suppose that the norm of X is Gâteaux differentiable at x. Then the mapping u_x is the unique normed linear functional u on X with $\|u\| = 1$ and $u(x) = \|x\|$.*

Proof. It is enough to take the case $\mathbb{K} = \mathbb{R}$. It is straightforward to show that u_x is convex, and that $u_x(y) = -u_x(-y)$ for each $y \in X$; whence, by Lemma 5.3.5, u_x is linear. Moreover, $|u_x(y)| \leqslant \|y\|$ and $u_x(x) = \|x\|$. So $\|u_x\| = 1$, and it only remains to prove the uniqueness. To this end, let u be any normed linear functional on X such that $\|u\| = 1$ and $u(x) = \|x\|$. For each $y \in X$ and each $t > 0$ we have

$$\frac{\|x - ty\| - \|x\|}{-t} \leqslant \frac{u(x - ty) - u(x)}{-t} = \frac{u(-ty)}{-t} = u(y)$$

and

$$u(y) = \frac{u(x + ty) - u(x)}{t} \leqslant \frac{\|x + ty\| - \|x\|}{t}.$$

Keeping y fixed and letting $t \longrightarrow 0$ in these two sets of displayed inequalities, we obtain $u_x(y) \leqslant u(y) \leqslant u_x(y)$; whence $u(y) = u_x(y)$. □

Proposition 5.3.7. *Let Y be a subspace of a uniformly convex Banach space X with Gâteaux differentiable norm, and let v be a nonzero normed linear functional on Y. Then there exists a unique normed linear functional u on X such that $\|u\| = \|v\|$ and $u(y) = v(y)$ for each $y \in Y$.*

Proof. We may assume that $\|v\| = 1$. Using Exercise 19 of Chapter 2 if necessary, we may further assume that Y is closed in X and is therefore a Banach space. Since Y is uniformly convex, it follows from Proposition 2.3.7 that there exists a unique $a \in Y$ such that $\|a\| = 1 = v(a)$. By Proposition 5.3.6, u_a is a normed linear functional on X such that $\|u_a\| = 1 = u_a(a)$. Applying the uniqueness part of Proposition 5.3.6 in the space Y, we conclude that $u_a(y) = v(y)$ for all $y \in Y$.

Finally, if u is a normed linear functional on X such that $\|u\| = \|v\|$ and $u(y) = v(y)$ for all $y \in Y$, then

$$u(a) = v(a) = 1 = \|v\| = \|u\|,$$

so, again by the uniqueness part of Proposition 5.3.6, $u = u_a$. $\square$

Not surprisingly, there is also a strong version of the separation theorem that applies in a uniformly convex separable Banach space; see Exercises 9 and 10.

5.4 Locally Convex Spaces

Although Errett Bishop considered that

> in most cases of interest it seems to be unnecessary to make use of any deep facts from the general theory of locally convex spaces ([9], page 350),

the development of constructive analysis (in particular, aspects of the theory of operators) in recent years has greatly benefited from such a general theory, which we now present.

A *locally convex space* consists of a linear space X over $\mathbb{K}$, a family $(p_i)_{i \in I}$ of seminorms on X, and the equality and (compatible, tight) inequality relations defined by

$$\begin{aligned} x = y &\iff \forall i \in I \ \ (p_i(x-y) = 0), \\ x \neq y &\iff \exists i \in I \ \ (p_i(x-y) > 0). \end{aligned}$$

Following normal practice, we call X itself a locally convex space when it is clear which family of seminorms is under consideration. The family $(p_i)_{i \in I}$ and the associated equality and inequality together form the *locally convex structure* on X. The corresponding *locally convex topology* on X is the family τ_X of all subsets of X that are unions of sets of the form

$$U(a, F, \varepsilon) = \left\{ x \in X : \sum_{i \in F} p_i(x-a) < \varepsilon \right\}$$

where $a \in X$, F is an inhabited finitely enumerable subset of I, and $\varepsilon > 0$. The seminorms p_i $(i \in I)$ are called the *defining seminorms* of τ_X. The members of τ_X are called *open subsets* of X, and the sets $U(a, F, \varepsilon)$ *basic neighbourhoods of* a. On the other hand, if S is a subset of X, then its *closure* $\overline{S}$ is the set of all elements x of X such that $S \cap U(x, F, \varepsilon)$ is inhabited for all finitely enumerable $F \subset I$ and all $\varepsilon > 0$. We say that S is *closed* (in the locally convex topology τ_X) if $S = \overline{S}$, and that S is *dense* in X if $\overline{S} = X$. The *unit ball* of the locally convex space X is

$$\{x \in X : \forall i \in I \ \ (p_i(x) \leqslant 1)\},$$

which is a closed subset of X.

We regard a normed space $(X, \| \ \|)$ as a locally convex space in which the family of seminorms consists of the single norm $\| \ \|$. In that case, the locally convex topology is just the metric topology associated with the norm on X, and the various notions of convergence, total boundedness, locatedness, and so on arising from the locally convex structure on X are precisely those associated with the norm on X.

In the following, unless we state otherwise, $\left(X, (p_i)_{i \in I}\right)$ and $\left(Y, (q_j)_{j \in J}\right)$ are locally convex spaces over $\mathbb{K}$. A mapping f of a subset S of X into Y is

- ▷ *continuous* at $a \in S$ if for each $\varepsilon > 0$ and each finitely enumerable $G \subset J$, there exist $\delta > 0$ and a finitely enumerable $F \subset I$ such that if $x \in S$ and $\sum_{i \in F} p_i(x - a) < \delta$, then $\sum_{j \in G} q_j(f(x) - f(a)) < \varepsilon$;
- ▷ *continuous on* S if it is continuous at each point of X;
- ▷ *uniformly continuous on* S if for each $\varepsilon > 0$ and each finitely enumerable subset G of J, there exist $\delta > 0$ and a finitely enumerable subset F of I such that if $x, y \in S$ and $\sum_{i \in F} p_i(x - y) < \delta$, then $\sum_{j \in G} q_j(f(x) - f(y)) < \varepsilon$.

Notice that each of the defining seminorms p_i on X is uniformly continuous on X.

Proposition 5.4.1. *The following are equivalent conditions on a linear mapping u between the locally convex spaces X and Y:*

(a) *u is continuous at* 0.

(b) *u is continuous on X.*

(c) *u is uniformly continuous on X.*

(d) *There exist a positive real number C and a finitely enumerable subset F of I such that*

$$|u(x)| \leqslant C \sup_{i \in F} p_i(x)$$

for each $x \in X$.

Proof. It is routine to show that (d) $\Longrightarrow$ (c) $\Longrightarrow$ (b) $\Longrightarrow$ (a). Suppose that u is continuous at 0. There exist $\delta > 0$ and a finitely enumerable subset F of I such that if $\sum_{i \in F} p_i(x) < \delta$, then $|u(x)| < 1$. It follows that for each $x \in X$ and each $\varepsilon > 0$,

$$\left| u \left(\frac{\delta x}{\sum_{i \in F} p_i(x) + \varepsilon} \right) \right| < 1$$

and therefore

$$|u(x)| < \frac{1}{\delta}\left(\sum_{i\in F} p_i(x) + \varepsilon\right).$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$|u(x)| \leqslant \frac{1}{\delta}\sum_{i\in F} p_i(x).$$

Thus (a) $\Longrightarrow$ (d). □

Naturally, in view of the analogous situation in a normed space, we say that a mapping u satisfying one, and hence all, of conditions (a)–(d) of Proposition 5.4.1 is a *bounded linear mapping.*

We say that a sequence $(x_n)_{n\geqslant 1}$ in a locally convex space $\left(X, (p_i)_{i\in I}\right)$

- *converges* to a *limit* x in X if for each basic neighbourhood $U = U(x, F, \varepsilon)$ of x in X there exists n_0 such that $x_n \in U$ whenever $n \geqslant n_0$;
- is a *Cauchy sequence* if for each $\varepsilon > 0$ and each finitely enumerable subset F of I, there exists n_0 such that $\sum_{i\in F} p_i(x_m - x_n) < \varepsilon$ whenever $m, n \geqslant n_0$.

We say that X is *complete* if every Cauchy sequence in X converges to a limit in X. (Strictly speaking, this definition applies only when X is *separable*—that is, has a countable dense subset. However, it will suffice for our purposes.)

Let S be a subset of X, F a finitely enumerable subset of I, and $\varepsilon > 0$. By an *ε-approximation to S relative to F* we mean a subset T of S such that for each $x \in X$ there exists $y \in T$ with $\sum_{i\in F} p_i\,(x - y) < \varepsilon$. We say that S is

- *totally bounded relative to F* if for each $\varepsilon > 0$ there exists a finitely enumerable ε-approximation to S relative to F;
- *totally bounded* if it is totally bounded relative to each finitely enumerable subset of I.

The proofs of the next five results are similar to those of their counterparts in metric space theory (see Chapter 2, Section 2) and are left as exercises.

Proposition 5.4.2. *If S is a totally bounded subset of a locally convex space, and f is a uniformly continuous mapping of S into a locally convex space Y, then $f(S)$ is totally bounded.*

Proposition 5.4.3. *If S is a totally bounded subset of a locally convex space, and f is a uniformly continuous mapping of S into $\mathbb{R}$, then $\sup_{x\in S} f(x)$ and $\inf_{x\in S} f(x)$ exist.*

A subset S of the locally convex space $\left(X,(p_i)_{i\in I}\right)$ is said to be *located* (in X) if

$$\inf\left\{\sum_{i\in F} p_i\,(x-y) : y\in S\right\}$$

exists for each $x\in X$ and each finitely enumerable subset F of I. As it does in metric spaces, locatedness plays an important role in locally convex spaces.

Proposition 5.4.4. *A totally bounded subset of a locally convex space is located.*

Proposition 5.4.5. *A located subset of a totally bounded set in a locally convex space is totally bounded.*

Theorem 5.4.6. *Let S be a totally bounded subset of a locally convex space X, and let f be a uniformly continuous mapping of S into $\mathbb{R}$. Then for all but countably many $t>\inf_{x\in S} f(x)$, the set*

$$\{x\in S : f(x)\geqslant t\}$$

is totally bounded.

Now consider the set $\mathcal{B}(X,Y)$ of all bounded linear mappings between the locally convex spaces X and Y. This set, which we often denote by $\mathcal{B}$ when it is clear which spaces X and Y are under consideration, becomes a locally convex space when taken with pointwise addition and multiplication-by-scalars and endowed with the seminorms p_x defined by

$$p_x(T)=\|Tx\| \quad (x\in X,\ T\in\mathcal{B}(X,Y))\,.$$

We denote the unit ball of $\mathcal{B}(X,Y)$ by $\mathcal{B}_1\,(X,Y)$ or just $\mathcal{B}_1$. When $X=Y$, we usually write $\mathcal{B}(X)$ and $\mathcal{B}_1(X)$ rather than $\mathcal{B}(X,Y)$ and $\mathcal{B}_1(X,Y)$.

In the special case where Y is the ground field $\mathbb{K}$, we obtain the space of all bounded linear functionals on X; this space is called the *dual space* of X, and is denoted by X^* or sometimes, for clarity, by $\left(X,(p_i)_{i\in I}\right)^*$; its unit ball is denoted by X_1^*. The topology associated with the family of seminorms $(p_x)_{x\in X}$ on X^* is called the *weak* topology* on X^*. When we are dealing with, for example, total boundedness relative to the locally convex structure on X^*, we speak of *weak*-total boundedness.*

In classical analysis, when X and Y are normed spaces, a big role is played by the *operator norm*

$$\|T\|=\sup\{\|Tx\| : x\in X,\ \ \|x\|\leqslant 1\}$$

of an element T of $\mathcal{B}(X,Y)$. Constructively the operator norm exists only for those elements T that are normed; but Corollary 4.1.8 shows that it exists for all $T\in$

$\mathcal{B}(X, Y)$ when X is finite-dimensional. In that case, X^* is algebraically isomorphic to X, every element of X^* is normed, the weak* topology on X^* coincides with the topology associated with the operator norm, and X^* is a Banach space relative to the operator norm; these facts are easily verified using results from Section 1 of Chapter 4.

We now examine the unit ball of the dual space.

Theorem 5.4.7. ***(Banach–Alaoglu theorem)*** *If X is a separable normed space, then X_1^* is complete and totally bounded relative to the locally convex structure on X^*.*

Proof. To prove the weak*-completeness, let $(v_n)_{n\geqslant 1}$ be a Cauchy sequence in X_1^*. Given $\varepsilon > 0$ and a finitely enumerable subset F of X, we can find N such that

$$\sum_{x\in F} |(v_m - v_n)(x)| < \varepsilon \quad (m, n \geqslant N)\,. \tag{5.7}$$

It readily follows that for each $x \in X$, $(v_n(x))_{n\geqslant 1}$ is a Cauchy sequence in $\mathbb{K}$. Since $\mathbb{K}$ is complete, the limit $v(x)$ of the sequence $(v_n(x))_{n\geqslant 1}$ exists in $\mathbb{K}$ for each $x \in X$. It is straightforward to show that the mapping $v : X \longrightarrow \mathbb{K}$ so defined is linear and belongs to X_1^*. Finally, keeping $n \geqslant N$ fixed and taking the limit with respect to m in (5.7), we obtain

$$\sum_{x\in F} |(v - v_n)\,(x)| \leqslant \varepsilon \qquad (n \geqslant N)\,.$$

Since ε, F are arbitrary, the sequence $(v_n)_{n\geqslant 1}$ converges to v in the weak* topology.

To prove the weak*-total boundedness of X_1^*, let $F = \{x_1, \ldots, x_m\}$ be a finitely enumerable subset of X, let

$$M > 4 + \max\{\|x_i\| : 1 \leqslant i \leqslant m\}\,,$$

and let $0 < \varepsilon < 1$. By Lemma 4.1.11, there exists a finite-dimensional subspace X_0 of X such that for $1 \leqslant i \leqslant m$, $\rho(x_i, X_0) < \varepsilon/m$ and therefore there exists $y_i \in X_0$ with $\|x_i - y_i\| < \varepsilon/m$. If $X_0 = \{0\}$, then for each $v \in X_1^*$,

$$\sum_{i=1}^{m} |v(x_i) - v(0)| \leqslant \sum_{i=1}^{m} \|x_i\| < \varepsilon,$$

so $\{0\}$ is an ε-approximation to X_1^* relative to F. We may therefore assume that X_0 has positive dimension. Then every element of X_0^* (the dual of X_0) is normed, and X_0^*, taken with the operator norm, is a finite-dimensional Banach space. Hence its unit ball is compact relative to the operator norm (Corollary 4.1.7). Moreover, by Proposition 2.3.6, each nonzero element of X_0^* has its kernel located in X_0; since X_0 is locally compact, this kernel is locally compact and hence is located *in the space* X (Proposition 2.2.18). It follows that the Hahn–Banach theorem can be applied to extend nonzero bounded linear functionals from X_0 to X.

Let $\{u_1^0, \ldots, u_n^0\}$ be an ε/m-approximation to the unit ball of X_0^* in the operator norm such that $0 < \|u_k^0\| < 1$ for each k. By the Hahn–Banach theorem, there exist normed linear functionals $u_1, \ldots, u_n$ in X_1^* such that $u_k(x) = u_k^0(x)$ for each $x \in X_0$ and each k. Given $u \in X_1^*$, since the restriction of u to X_0 is in the unit ball of X_0^*, we can find k such that $|u(x) - u_k^0(x)| < \varepsilon/m$ for all $x \in X_0$ with $\|x\| \leqslant 1$. For each i $(1 \leqslant i \leqslant m)$ we have

$$\begin{aligned}
|u(x_i) - u_k(x_i)| &\leqslant |(u - u_k)(x_i - y_i)| + \left|\left(u - u_k^0\right)(y_i)\right| \\
&\leqslant 2\|x_i - y_i\| + (1 + \|y_i\|)\left|\left(u - u_k^0\right)\left(\frac{1}{1 + \|y_i\|} y_i\right)\right| \\
&< \frac{2\varepsilon}{m} + \left(1 + \|x_i\| + \frac{\varepsilon}{m}\right)\frac{\varepsilon}{m} \\
&< \frac{M\varepsilon}{m}.
\end{aligned}$$

Hence

$$\sum_{i=1}^{m} |u(x_i) - u_k(x_i)| < M\varepsilon. \tag{5.8}$$

We now see that $\{u_1, \ldots, u_n\}$ is an $M\varepsilon$-approximation to X_1^* relative to F. □

Corollary 5.4.8. *If X is a separable normed space, then the normed linear functionals on X are weak*-dense in X^*.*

Proof. In the notation of the last paragraph of the foregoing proof, the function u_k satisfying (5.8) is normed. Since $\varepsilon > 0$ and the finitely enumerable subset F of X are arbitrary, the result follows. □

Let X be a normed space. It is easy to see that, for a fixed vector $x \in X$, the linear functional $u \rightsquigarrow u(x)$ on X^* is weak*-uniformly continuous on X_1^*. We shall show that any element of X^{**} (the dual of X^*) that is uniformly continuous on X_1^* can be approximated arbitrarily closely by functionals of this special form, and that if X is complete, then this approximation can be made exact.

Lemma 5.4.9. *Let V be a locally convex space whose unit ball V_1 is totally bounded, and let ϕ be a nonzero linear functional on V that is uniformly continuous on V_1. Then the **unit kernel**, $V_1 \cap \ker(\phi)$, of ϕ is totally bounded.*

Proof. Since ϕ is nonzero and uniformly continuous on the totally bounded set V_1,

$$C = \sup\{|\phi(y)| : y \in V_1\}$$

exists and is positive. Using the linearity of ϕ, choose $y \in V_1$ with $\phi(y) > C/2$. Then

$$y_0 = \frac{C}{2\phi(y)} y$$

belongs to V_1, and $\phi(y_0) = C/2$. Let ε be a positive number, and

$$F = \{p_1, \dots, p_m\}$$

a finitely enumerable set of defining seminorms on V. Using Theorem 5.4.6, compute a positive number

$$t < \frac{C\varepsilon}{C + 4m}$$

such that

$$S_t = \{y \in V_1 : |\phi(y)| \leqslant t\}$$

is totally bounded. Pick a t-approximation $\{s_1, \dots, s_n\}$ of S_t relative to F. Set

$$y_k = \frac{C}{C+2t}\left(s_k - \frac{2}{C}\phi(s_k) y_0\right) \quad (1 \leqslant k \leqslant n).$$

Then $y_k \in \ker(\phi)$. If p_i is any defining seminorm on V, then we have

$$\begin{aligned} p_i(y_k) &\leqslant \frac{C}{C+2t}\left(p_i(s_k) + \frac{2}{C}|\phi(s_k)| p_i(y_0)\right) \\ &\leqslant \frac{C}{C+2t}\left(1 + \frac{2t}{C}\right) \\ &= 1. \end{aligned}$$

So y_k belongs to $V_1 \cap \ker(\phi)$.

Now consider any element y of $V_1 \cap \ker(\phi)$. Since $y \in S_t$, there exists k such that $\sum_{i=1}^{m} p_i(y - s_k) < t$ and therefore

$$\begin{aligned} \sum_{i=1}^{m} p_i(y - y_k) &\leqslant \sum_{i=1}^{m} p_i(y - s_k) + \sum_{i=1}^{m} p_i(s_k - y_k) \\ &< t + \frac{2}{C+2t}\sum_{i=1}^{m} p_i(t s_k + \phi(s_k) y_0) \\ &\leqslant t + \frac{2t}{C}\sum_{i=1}^{m} (p_i(s_k) + p_i(y_0)) \\ &\leqslant t\left(1 + \frac{4m}{C}\right) \\ &< \varepsilon. \end{aligned}$$

Thus $\{y_1, \dots, y_n\}$ is a finitely enumerable ε-approximation to $V_1 \cap \ker(\phi)$. $\square$

Lemma 5.4.10. *Let X be a normed linear space, and ϕ a linear functional on X^* that is uniformly continuous on X_1^*. Then for each $\varepsilon > 0$ there exist $\delta > 0$ and a finite-dimensional subspace X_0 of X such that if $u \in X_1^*$ and $|u(x)| < \delta$ for all x in the unit ball of X_0, then $|\phi(u)| < \varepsilon$.*

Proof. There exist $\xi_1, \ldots, \xi_n$ in X with $\|\xi_k\| < 1/2$ for each k, and a positive number δ, such that if $u \in X_1^*$ and $|u(\xi_k)| < 2\delta$ for each k, then $|\phi(u)| < \varepsilon$. By Lemma 4.1.11, there exist a finite-dimensional subspace X_0 of X, and vectors $x_1, \ldots, x_n$ of X_0, such that

$$\|\xi_k - x_k\| < \min\left\{\frac{1}{2}, \delta\right\} \quad (1 \leqslant k \leqslant n).$$

Suppose that $u \in X_1^*$ and $|u(x)| < \delta$ for all x in the unit ball of X_0. Then for each k,

$$|u(\xi_k)| \leqslant |u(\xi_k - x_k)| + |u(x_k)| \leqslant \|\xi_k - x_k\| + \delta < 2\delta.$$

Hence $|\phi(u)| < \varepsilon$. □

Lemma 5.4.11. *Let X be a separable normed space, X_0 a finite-dimensional subspace of X, and ϕ a nonzero linear functional on X that is weak*-uniformly continuous on X_1^*. Denote the unit kernel of ϕ by $\mathcal{N}_1$. Then the following hold:*

(a) $\|\phi\| = \sup\{|\phi(u)| : u \in X_1^*\}$ *exists.*

(b) $\|x\|_0 = \sup\{|u(x)| : u \in \mathcal{N}_1\}$ *defines a seminorm on X_0.*

(c) $\beta = \inf\{\|x\|_0 : x \in X_0, \ \|x\| = 1\}$ *is well defined.*

Suppose that $\beta > 0$. Then

(d) $(X_0, \|\ \|_0)$ *is a finite-dimensional Banach space, and*

$$\|u\|^0 = \sup\{|(u(x)| : x \in X_0, \ \|x\|_0 \leqslant 1\}$$

defines the corresponding operator norm on the dual space X_0^*.

Finally, for each $u \in \mathcal{N}_1$ let $F(u)$ denote the restriction of u to X_0^*. Then

(e) *F is weak*-uniformly continuous, and* $\operatorname{ran}(F)$ *is $\|\ \|^0$-dense in the unit ball S_0^* of* $\left(X_0^*, \|\ \|^0\right)$.

Proof. Since ϕ is weak*-uniformly continuous on X_1^* and, by Theorem 5.4.7, the latter is weak*-totally bounded, we see that $\|\phi\|$ exists. Also, $\mathcal{N}_1$ is weak*-totally bounded, by Lemma 5.4.9, and the mapping $u \rightsquigarrow u(x)$ is weak*-uniformly continuous on X^*, so $\|x\|_0$ exists. Clearly, $\|\ \|_0$ is a seminorm and $\|\ \|_0 \leqslant \|\ \|$; whence $\|\ \|_0$ is uniformly continuous on $(X_0, \|\ \|)$. Since that space is finite-dimensional,

$$\{x \in X_0 : \|x\| = 1\}$$

is compact, and so β exists.

Now suppose that $\beta > 0$. Then $\| \ \| \leqslant \beta^{-1} \| \ \|_0$, so $\| \ \|_0$ is a norm on X_0 equivalent to $\| \ \|$, and $(X_0, \| \ \|_0)$ is a finite-dimensional Banach space. The mapping $F : \mathcal{N}_1 \longrightarrow (X_0^*, \| \ \|^0)$ is weak*-uniformly continuous on the weak*-totally bounded set $\mathcal{N}_1$, so its range is totally bounded, and hence located, in $(X_0^*, \| \ \|^0)$. To show that $\mathrm{ran}(F)$ is dense in the unit ball S_0^* of $(X_0^*, \| \ \|^0)$, fix u_0 in S_0^* and suppose that

$$0 < c = \inf \left\{ \|u_0 - F(u)\|^0 : u \in \mathcal{N}_1 \right\}.$$

By Corollary 5.2.10, there exists a normed linear functional Φ on $(X_0^*, \| \ \|^0)$ such that

$$\Phi(u_0) > |\Phi(F(u))| + \frac{c}{2} \quad (u \in \mathcal{N}_1).$$

Now X_0^*, being finite-dimensional, is equivalent to a Hilbert space. (Recall from Corollary 4.1.9 that all norms on a finite-dimensional space are equivalent to each other and hence to the Euclidean norm.) It follows from this and the Riesz representation theorem that there exists $\zeta \in X_0$ such that $\Phi(v) = v(\zeta)$ for each $v \in X_0^*$. Therefore

$$u_0(\zeta) > \sup\{|F(u)(\zeta)| : u \in \mathcal{N}_1\} = \sup\{|u(\zeta)| : u \in \mathcal{N}_1\} = \|\zeta\|_0 .$$

This is absurd, since $u_0 \in S_0^*$. Hence, in fact, $c = 0$ and $\mathrm{ran}(F)$ is $\| \ \|^0$-dense in S_0^*. □

Lemma 5.4.12. *With the hypotheses and notation of* Lemma 5.4.11 *and its proof, suppose that $\|\phi\| > 1$ and let α, δ be positive numbers. Suppose also that $|\phi(u)| < \alpha/2$ whenever* (a) *$u \in X_1^*$ and* (b) *$|u(x)| < \delta$ for all x in the unit ball of X_0. Then $\beta < \alpha$.*

Proof. Choose $u_0 \in X_1^*$ such that $\phi(u_0) = 1$. Since

$$|u_0(\beta x)| = \beta |u_0(x)| \leqslant \beta \|x\| \leqslant \|x\|_0 ,$$

the linear functional

$$\Psi : x \rightsquigarrow u_0(\beta x)$$

belongs to S_0^*. Now, either $\beta < \alpha$ or $\beta > 0$. Assuming the latter inequality, we see from Lemma 5.4.11 that there exists $u \in \mathcal{N}_1$ such that

$$\|\Psi - F(u)\|^0 < 2\delta.$$

So for all x in the unit ball of $(X_0, \| \ \|_0)$,

$$|u_0(\beta x) - u(x)| < 2\delta$$

and therefore

$$\left|\frac{1}{2}\left(\beta u_0 - u\right)(x)\right| < \delta.$$

But $\beta \leqslant 1$, and both u_0 and $-u$ belong to the convex set X_1^*, so $\frac{1}{2}\left(\beta u_0 - u\right) \in X_1^*$; whence

$$\left|\phi\left(\frac{1}{2}\left(\beta u_0 - u\right)\right)\right| < \frac{\alpha}{2}.$$

Therefore

$$\beta = \beta\phi(u_0) - \phi(u) = 2\phi\left(\frac{1}{2}\left(\beta u_0 - u\right)\right) < \alpha,$$

as we required. □

Proposition 5.4.13. *Let X be a separable normed space, and ϕ a linear functional on X^* that is weak*-uniformly continuous on X_1^*. Then for each $\varepsilon > 0$ there exists $x \in X$ such that $\|x\| < 3\|\phi\|$ and*

$$|\phi(u) - u(x)| < \varepsilon \qquad \left(u \in X_1^*\right).$$

Proof. If $\|\phi\| < \varepsilon$, then we can take $x = 0$; so we may assume that $\|\phi\| > 0$. Scaling if necessary, we may also assume that there exists $u_0 \in X_1^*$ with $\phi(u_0) = 1$. Pick a positive number α, which we shall specify further as the proof develops. Using Lemma 5.4.10, construct a finite-dimensional subspace X_0 of X, and a positive number δ, such that if $u \in X_1^*$ and $|u(x)| < \delta$ for all x in the unit ball of X_0, then $|\phi(u)| < \alpha/2$. Applying Lemma 5.4.12, we see that there exists $z \in X_0$ such that $\|z\| = 1$ and $|u(z)| < \alpha$ for all $u \in \mathcal{N}_1$. For each $u \in X_1^*$, since

$$\frac{1}{1 + \|\phi\|}\left(u - \phi(u)u_0\right) \in \mathcal{N}_1,$$

we therefore have

$$\frac{1}{1 + \|\phi\|}\left|\left(u - \phi(u)u_0\right)(z)\right| < \alpha.$$

By Proposition 5.3.1, there exists $u_1 \in X_1^*$ such that $u_1(z) > 1/2$. Thus

$$\frac{1}{2} < \left|\left(u_1 - \phi(u_1)u_0\right)(z)\right| + |\phi(u_1)|\,|u_0(z)| \leqslant \left(1 + \|\phi\|\right)\alpha + \|\phi\|\,|u_0(z)|,$$

and therefore

$$|u_0(z)| > \frac{1}{\|\phi\|}\left(\frac{1}{2} - \left(1 + \|\phi\|\right)\alpha\right) > \frac{1}{3\|\phi\|}$$

provided we choose α small enough. In that case, writing

$$x = \frac{1}{u_0(z)}z,$$

we see that $\|x\| < 3\|\phi\|$. Also, for each $u \in X_1^*$,

$$|u(x) - \phi(u)| = \frac{1}{|u_0(z)|} |(u - \phi(u)u_0)(z)| < 3\|\phi\|(1 + \|\phi\|)\alpha,$$

which can be made less than ε by a suitably small choice of α. □

Theorem 5.4.14. *Let X be a separable Banach space, and ϕ a linear functional on X^* that is weak*-uniformly continuous on X_1^*. Then there exists $x \in X$ such that $\phi(u) = u(x)$ for each $u \in X^*$.*

Proof. We may assume that $\|\phi\| < 1$. Recursively applying Proposition 5.4.13, construct a sequence $(x_n)_{n \geqslant 1}$ of vectors in X such that for each n,

$$\left| \phi(u) - \sum_{k=1}^{n} u(x_k) \right| < \frac{1}{2^n} \qquad (u \in X_1^*) \tag{5.9}$$

and $\|x_n\| < 3/2^{n-1}$. The series $\sum_{n=1}^{\infty} x_n$ then converges to an element x of the complete space X. Using the linearity and continuity of u, and letting $n \longrightarrow \infty$ in (5.9), we obtain the desired conclusion. □

Let H be a nontrivial Hilbert space. One of the topologies on $\mathcal{B}(H)$ that is important in operator-algebra theory is the *weak-operator topology* τ_w*:* that is, the locally convex topology defined by the seminorms of the form $T \rightsquigarrow |\langle Tx, y\rangle|$ with x, y in H. Classically—but *not* constructively (see Exercise 17)—the sets of the type

$$\left\{ T \in \mathcal{B}(H) : \sum_{i,j=1}^{n} |\langle Te_i, e_j\rangle| < \delta \right\},$$

with $\delta > 0$ and $\{e_1, \ldots, e_n\}$ a set of pairwise orthogonal unit vectors in H, form a base of weak-operator neighbourhoods of 0 in $\mathcal{B}(H)$, so a linear functional ϕ on $\mathcal{B}(H)$ is τ_w-continuous if and only if it has the following special continuity property:

SC *There exist $\delta > 0$ and a set $\{e_1, \ldots, e_n\}$ of pairwise orthogonal unit vectors in H such that for each $T \in \mathcal{B}(H)$, if $\sum_{i,j=1}^{n} |\langle Te_i, e_j\rangle| < \delta$, then $|\phi(T)| < 1$.*

We shall use the technique embodied in the proofs of Lemma 5.4.11, Lemma 5.4.12, and Proposition 5.4.13 to produce a characterisation of those linear functionals ϕ on $\mathcal{B}(H)$ with the property **SC**.

Before the next result, we mention that, in spite of what we wrote in the preceding paragraph, it is constructively provable that the sets of the form

$$\left\{ T \in \mathcal{B}_1(H) : \sum_{i,j=1}^{n} |\langle Te_i, e_j \rangle| < \delta \right\},$$

with $\delta > 0$ and $\{e_1, \ldots, e_n\}$ a set of pairwise orthogonal unit vectors in H, form a base of weak-operator neighbourhoods of 0 *in the unit ball* $\mathcal{B}_1(H)$. This is left as part of Exercise 17.

Proposition 5.4.15. *If H is a nontrivial Hilbert space, then $\mathcal{B}_1(H)$ is τ_w-totally bounded.*

Proof. Let $\{e_1, \ldots, e_n\}$ be a finite set of pairwise orthogonal unit vectors generating a finite-dimensional subspace H_0 of H. In view of the observation preceding this proposition, it will suffice to prove that $\mathcal{B}_1(H)$ is totally bounded with respect to the seminorm $p_{jk} : T \rightsquigarrow \sum_{j,k=1}^{n} |\langle Te_j, e_k \rangle|$. Let P be the projection of H on H_0. Note that $\mathcal{B}(H_0)$ is a finite-dimensional Banach space, and hence has a totally bounded unit ball, relative to the operator norm. Let $\{T_1^0, \ldots, T_m^0\}$ be an ε/n^2-approximation to $\mathcal{B}_1(H_0)$, and consider any $T \in \mathcal{B}_1(H)$. The restriction $(PT)_0$ of PT to H_0 belongs to $\mathcal{B}_1(H_0)$, so there exists i such that $\left\|(PT)_0 - T_i^0\right\| < \varepsilon/n^2$. Also, $T_i^0 P \in \mathcal{B}_1(H)$. Thus if $1 \leqslant j, k \leqslant n$, then

$$\begin{aligned} \left|\left\langle \left(T - T_i^0 P\right) e_j, e_k \right\rangle\right| &= \left|\left\langle \left(T - T_i^0\right) e_j, P e_k \right\rangle\right| = \left|\left\langle P\left(T - T_i^0\right) e_j, e_k \right\rangle\right| \\ &= \left|\left\langle \left((PT)_0 - T_i^0\right) e_j, e_k \right\rangle\right| \leqslant \left\|(PT)_0 - T_i^0\right\| < \frac{\varepsilon}{n^2}. \end{aligned}$$

Hence $\sum_{j,k=1}^{n} \left|\left\langle \left(T - T_i^0 P\right) e_j, e_k \right\rangle\right| < \varepsilon$. We now see that $\{T_1^0 P, \ldots, T_m^0 P\}$ is an ε-approximation to $\mathcal{B}_1(H)$ relative to the seminorm p_{jk}. □

Proposition 5.4.16. *Let H be a nontrivial Hilbert space, and let ϕ be a linear functional on $\mathcal{B}(H)$ with the property* ***SC****. Then for each $\varepsilon > 0$ there exist a finite set $\{e_1, \ldots, e_n\}$ of pairwise orthogonal unit vectors in H, and elements c_{jk} $(1 \leqslant j, k \leqslant n)$ of $\mathbb{K}$, such that*

$$\left| \phi(T) - \sum_{j,k=1}^{n} c_{jk} \langle Te_j, e_k \rangle \right| < \varepsilon \quad (T \in \mathcal{B}_1(H)).$$

Proof. With the proofs of Lemma 5.4.11, Lemma 5.4.12, and Proposition 5.4.13 at hand, we omit some of the grisly details of this one. We first note that, in view of the observation immediately preceding Proposition 5.4.15, ϕ is τ_w-uniformly continuous on $\mathcal{B}_1(H)$. It follows from this, Proposition 5.4.15, and Proposition 5.4.3 that

$$\|\phi\| = \sup\{|\phi(T)| : T \in \mathcal{B}_1(H)\}$$

exists. Scaling if necessary, we may assume that there exists $T_0 \in \mathcal{B}_1(H)$ such that $\phi(T_0) = 1$. Choose a finite set $\{e_1, \ldots, e_n\}$ of pairwise orthogonal unit vectors in H, and a positive number δ, such that if $T \in \mathcal{B}(H)$ and $\sum_{i,j=1}^n |\langle Te_i, e_j\rangle| < 2\delta$, then $|\phi(T)| < 1$. Let $H_0 = \operatorname{span}\{e_1, \ldots, e_n\}$, and for each $\mathbf{x} = (x_1, \ldots, x_n)$ in the n-fold product space H_0^n, define the product norm

$$\|\mathbf{x}\| = \sum_{k=1}^n \|x_k\|.$$

Note that for each $T \in \mathcal{B}_1(H)$ and each $\mathbf{x} \in H_0^n$,

$$\left|\sum_{k=1}^n \langle Tx_k, e_k\rangle\right| \leqslant \sum_{k=1}^n |\langle Tx_k, e_k\rangle| \leqslant \sum_{k=1}^n \|x_k\| = \|\mathbf{x}\|. \tag{5.10}$$

For such $\mathbf{x}$, the mapping

$$T \rightsquigarrow \left|\sum_{k=1}^n \langle Tx_k, e_k\rangle\right|$$

is τ_w-uniformly continuous on $\mathcal{B}_1(H)$. Since (by Proposition 5.4.15 and Lemma 5.4.9) the unit kernel $\mathcal{N}_1$ of ϕ is τ_w-totally bounded, it follows from Proposition 5.4.3 that we can define a seminorm on H_0^n by

$$\|\mathbf{x}\|_0 = \sup\left\{\left|\sum_{k=1}^n \langle Tx_k, e_k\rangle\right| : T \in \mathcal{N}_1\right\}.$$

Then $\|\ \|_0 \leqslant \|\ \|$, by (5.10), so the mapping $\|\ \|_0$ from $(H_0^n, \|\ \|)$ to $\mathbb{R}$ is uniformly continuous. Since H_0^n is finite-dimensional, the set

$$\{\mathbf{x} \in H_0^n : \|\mathbf{x}\| = 1\}$$

is $\|\ \|$-compact. Hence

$$\beta = \inf\{\|\mathbf{x}\|_0 : \mathbf{x} \in H_0^n,\ \|\mathbf{x}\| = 1\}$$

is well defined. This time (cf. Lemma 5.4.12) we prove that $\beta = 0$. Suppose that $\beta > 0$. Then $(H_0^n, \|\ \|_0)$ is a finite-dimensional Banach space, and

$$\|u\|^0 = \sup\{|u(\mathbf{x})| : \mathbf{x} \in H_0^n,\ \|\mathbf{x}\|_0 \leqslant 1\}$$

defines the corresponding operator norm on the dual space $(H_0^n)^*$. For each $T \in \mathcal{N}_1$ and each $\mathbf{x} \in H_0^n$ let

$$F(T)(\mathbf{x}) = \sum_{k=1}^{n} \langle Tx_k, e_k \rangle .$$

The mapping $F : \mathcal{N}_1 \longrightarrow ((H_0^n)^*, \| \|^0)$ is τ_w-uniformly continuous, so $F(\mathcal{N}_1)$ is totally bounded, and therefore located, in $((H_0^n)^*, \| \|^0)$. Using the separation theorem as in the proof of Lemma 5.4.12, we can show that $F(\mathcal{N}_1)$ is dense in the unit ball S_0^* of $(H_0^n)^*, \| \|^0)$. Since the linear functional

$$\Psi : \mathbf{x} \rightsquigarrow \sum_{k=1}^{n} \langle \beta T_0 x_k, e_k \rangle$$

belongs to S_0^*, given $t > 0$ we can find $T \in \mathcal{N}_1$ such that $\|\Psi - F(T)\|^0 < n^{-2}\delta t$. Write

$$\{(j,k) : 1 \leqslant j,k \leqslant n\} = P \cup Q,$$

where P and Q are disjoint sets,

$$(j,k) \in P \Longrightarrow \langle (\beta T_0 - T)\, e_j, e_k \rangle \neq 0,$$
$$(j,k) \in Q \Longrightarrow |\langle (\beta T_0 - T)\, e_j, e_k \rangle| < n^{-2}\delta t.$$

If $(j,k) \in P$, set

$$r_{jk} = \langle (\beta T_0 - T)\, e_j, e_k \rangle^{-1} |\langle (\beta T_0 - T)\, e_j, e_k \rangle| ,$$

and if $(j,k) \in Q$, set $r_{jk} = 0$. Now define $\boldsymbol{\xi} \in H_0^n$ by

$$\xi_k = n^{-2} \sum_{j=1}^{n} r_{jk} e_j \quad (1 \leqslant k \leqslant n) ,$$

to obtain

$$\|\xi_k\| \leqslant n^{-2} \sum_{j=1}^{n} |r_{jk}| \leqslant n^{-1}$$

and hence

$$\|\boldsymbol{\xi}\|_0 \leqslant \|\boldsymbol{\xi}\| = \sum_{k=1}^{n} \|\xi_k\| \leqslant 1.$$

Thus

$$\left| \sum_{k=1}^{n} \langle (\beta T_0 - T)\, \xi_k, e_k \rangle \right| < n^{-2}\delta t. \tag{5.11}$$

Moreover,

$$
\begin{aligned}
n^2\left|\sum_{k=1}^{n}\langle(\beta T_0-T)\,\xi_k,e_k\rangle\right| &= \left|\sum_{k=1}^{n}\sum_{j=1}^{n} r_{jk}\,\langle(\beta T_0-T)\,e_j,e_k\rangle\right| \\
&= \left|\sum_{(j,k)\in P} r_{jk}\,\langle(\beta T_0-T)\,e_j,e_k\rangle\right| \\
&= \sum_{(j,k)\in P} \left|\langle(\beta T_0-T)\,e_j,e_k\rangle\right|,
\end{aligned}
$$

so

$$
\begin{aligned}
\sum_{j,k=1}^{n}\left|\langle(\beta T_0-T)\,e_j,e_k\rangle\right| &\leqslant \sum_{(j,k)\in P}\left|\langle(\beta T_0-T)\,e_j,e_k\rangle\right| + \sum_{(j,k)\in Q}\left|\langle(\beta T_0-T)\,e_j,e_k\rangle\right| \\
&\leqslant n^2\left|\sum_{k=1}^{n}\langle(\beta T_0-T)\,\xi_k,e_k\rangle\right| + n^2\left(n^{-2}\delta t\right) \leqslant 2\delta t,
\end{aligned}
$$

the last step following from (5.11). Hence

$$
\sum_{j,k=1}^{n}\left|t^{-1}\,\langle(\beta T_0-T)\,e_j,e_k\rangle\right| < 2\delta
$$

and therefore $|\phi(\beta T_0-T)|<t$. We now have

$$
\beta=\beta\phi(T_0)-\phi(T)=\phi(\beta T_0-T)<t.
$$

Since $t>0$ is arbitrary, it follows that $\beta=0$.

Thus for each $\alpha>0$ there exists $\mathbf{z}\in H_0^n$ such that $\|\mathbf{z}\|=1$ and

$$
\left|\sum_{k=1}^{n}\langle Tz_k,e_k\rangle\right| < (2n)^{-1}\alpha \quad (T\in\mathcal{N}_1)\,.
$$

We can arrange that each $z_k\neq 0$, provided we replace $(2n)^{-1}$ by n^{-1} in the last inequality; we assume that this has been done. Therefore

$$
\left|\sum_{k=1}^{n}\langle(T-\phi(T)T_0)\,z_k,e_k\rangle\right| < n^{-1}\left(1+\|\phi\|\right)\alpha \quad (T\in\mathcal{B}_1(H))\,.
$$

We now introduce an operator T_1 that will enable us to bound $\sum_{k=1}^{n}\langle T_0z_k,e_k\rangle$ away from 0. Setting

$$
T_1x=n^{-1}\sum_{k=1}^{n}\|z_k\|^{-1}\langle x,z_k\rangle\,e_k \quad (x\in H)\,,
$$

we have

$$\|T_1 x\| \leqslant n^{-1} \sum_{k=1}^{n} \|z_k\|^{-1} \|x\| \|z_k\| = \|x\|,$$

so $T_1 \in \mathcal{B}_1(H)$. Also,

$$\sum_{k=1}^{n} \langle T_1 z_k, e_k \rangle = n^{-1} \sum_{k=1}^{n} \|z_k\|^{-1} \langle z_k, z_k \rangle = n^{-1} \sum_{k=1}^{n} \|z_k\| = n^{-1}.$$

Hence

$$\begin{aligned} n^{-1} &\leqslant \left| \sum_{k=1}^{n} \langle (T_1 - \phi(T_1) T_0) z_k, e_k \rangle \right| + |\phi(T_1)| \left| \sum_{k=1}^{n} \langle T_0 z_k, e_k \rangle \right| \\ &\leqslant n^{-1} (1 + \|\phi\|) \alpha + \|\phi\| \left| \sum_{k=1}^{n} \langle T_0 z_k, e_k \rangle \right| \end{aligned}$$

and therefore

$$\left| \sum_{k=1}^{n} \langle T_0 z_k, e_k \rangle \right| \geqslant (n \|\phi\|)^{-1} (1 - (1 + \|\phi\|) \alpha) > (2n \|\phi\|)^{-1}$$

provided α is small enough. Setting

$$\mathbf{x} = \left(\sum_{k=1}^{n} \langle T_0 z_k, e_k \rangle \right)^{-1} \mathbf{z},$$

for each $T \in \mathcal{B}_1(H)$ we now obtain

$$\begin{aligned} \left| \phi(T) - \sum_{k=1}^{n} \langle T x_k, e_k \rangle \right| &= \left| \sum_{k=1}^{n} \langle T_0 z_k, e_k \rangle \right|^{-1} \left| \sum_{k=1}^{n} \langle \phi(T) T_0 z_k, e_k \rangle - \sum_{k=1}^{n} \langle T z_k, e_k \rangle \right| \\ &< 2n \|\phi\| n^{-1} (1 + \|\phi\|) \alpha \\ &= 2 \|\phi\| (1 + \|\phi\|) \alpha, \end{aligned}$$

which can be made less than ε by choosing α suitably small. It remains to take $c_{jk} = \langle x_k, e_j \rangle$ for $1 \leqslant j, k \leqslant n$. □

Exercises

1. Let U, V be subsets of a Banach space X such that $U \cup V$ is dense. Prove the following:
   (a) If $u \in U$ and $v \in V$, then for each $\varepsilon > 0$ there exist $t \in [0, 1]$ and $x \in \overline{U} \cap \overline{V}$ such that $\|x - tu - (1 - t)v\| < \varepsilon$.

(b) For each $x \in X$,

$$\rho\left(x, \overline{U} \cap \overline{V}\right) = \min\left\{\rho(x, U), \rho(x, V)\right\},$$

in the sense that each side of this equation exists if and only if the other does.

2. We say that a subset S of a normed space X has the *boundary crossing property* if for each $\xi \in S$, each $z \in \sim S$, and each $\varepsilon > 0$ there exist $t \in [0,1]$ and $x \in \partial S$ such that $\|x - t\xi - (1-t)\,z\| < \varepsilon$. Prove that if X is complete and $S \cup \sim S$ is dense, then S has the boundary crossing property. Do we necessarily obtain this conclusion if $S \cup \sim S$ is not known to be dense in X?

3. Let S be a subset of a Banach space such that $S \cup \sim S$ is dense, and let $x_0 \in \sim S$. Prove that $\rho\,(x_0, \partial S)$ exists if and only if $\rho\,(x_0, S)$ exists, in which case these two distances are equal.

4. Let $\{e_1, \ldots, e_n\}$ be the canonical orthonormal basis of the Euclidean space $\mathbb{R}^n$, and Σ the simplex whose vertices are $\pm e_j$ $(1 \leqslant j \leqslant n)$. Let ξ belong to the metric complement $-\Sigma$ of Σ in $\mathbb{R}^n$. Prove that the segment $[0, \xi]$ contains points arbitrarily close to the union of the faces of Σ.

5. Show that if the Hahn–Banach theorem holds with exact preservation of the norm of the functional, then LLPO holds.

6. Let Y be a finite-dimensional subspace of a normed space X, and $x_0 \in X$. Suppose that $\rho(x_0, Y) > 0$ and that X is the span of $Y \cup \{x_0\}$. Let v be a nonzero normed linear functional on Y, and $\varepsilon > 0$. Without using the separation theorem, the Hahn–Banach theorem, or any of their consequences, prove that there exists a normed linear functional u on X such that $\|u\| \leqslant \|v\| + \varepsilon$ and $u(y) = v(y)$ for each $y \in Y$. Use this to give another proof of the Hahn–Banach theorem.

7. Use the Hahn–Banach theorem to prove Proposition 5.3.1.

8. Let p be an integer $\geqslant 2$. Prove that the Banach space l_p is uniformly convex and has Gâteaux differentiable norm. (See Exercise 11 of Chapter 4.)

9. Let X be a normed space with Gâteaux differentiable norm. Let C be a closed, located, convex subset of X, let $x \in -C$, and let $y \in C$. Prove that $\|x - y\| = \rho\,(x, C)$ if and only if $\operatorname{Re} u_{x-y}\,(z - y) \leqslant 0$ for all $z \in C$.

10. Let X be a uniformly convex Banach space with Gâteaux differentiable norm. Let A and B be subsets of X whose algebraic difference is located and convex, and whose mutual distance is positive. Prove that there exists a normed linear functional u on X, with norm 1, such that $\operatorname{Re} u\,(y) \geqslant \operatorname{Re} u\,(x) + d$ for all $x \in A$

and $y \in B$. (The results embodied in Exercises 9 and 10 are due to Ishihara [55].)

11. Prove that if, under the other hypotheses of Lemma 5.4.9, the conclusion holds without the hypothesis that ϕ is either 0 or nonzero, then we can derive the essentially nonconstructive proposition

$$\forall x \in \mathbb{R} \; (x = 0 \vee \neg (x = 0)).$$

12. Let f be a uniformly continuous linear functional on X, and S a balanced, totally bounded subset of X. Prove that the set

$$S_t = \{x \in S : |f(x)| \leqslant t\}$$

is totally bounded for each $t > 0$.

13. Let (X, p) be a seminormed space, and S a balanced, totally bounded subset of X. Let the mapping $f : X \longrightarrow \mathbb{K}$ be both uniformly continuous on S and homogeneous—that is, $f(\lambda x) = \lambda f(x)$ for all $\lambda \in \mathbb{K}$ and $x \in X$. Prove that the set

$$S_t = \{x \in S : |f(x)| \leqslant t\}$$

is totally bounded for each $t > 0$.

14. Let X be a finite-dimensional locally convex space, and let S be the set of convex combinations of points in an inhabited finitely enumerable subset of X. Prove that S is totally bounded in X.

15. Fill in the missing details in the proof of Proposition 5.4.16.

16. Let $(e_n)_{n \geqslant 1}$ be an orthonormal basis of a separable Hilbert space H, and let ϕ be a linear functional on $\mathcal{B}(H)$ that is weak-operator uniformly continuous on $\mathcal{B}_1(H)$. Prove that there exist complex numbers c_{jk} $(j, k \geqslant 1)$ such that $\phi(T) = \sum\limits_{j,k=1}^{\infty} c_{jk} \langle Te_j, e_k \rangle$ for all $T \in \mathcal{B}(H)$.

17. Prove the following for a nontrivial Hilbert space H:

    (a) The sets of the form

$$\left\{ T \in \mathcal{B}_1(H) : \sum_{i,j=1}^{n} |\langle Te_i, e_j \rangle| < \delta \right\},$$

    with $\delta > 0$ and $\{e_1, \ldots, e_n\}$ a set of pairwise orthogonal unit vectors in H, form a base of weak-operator neighbourhoods of 0 in the unit ball $\mathcal{B}_1(H)$.

(b) If $\mathcal{B}_1(H)$ can be replaced by $\mathcal{B}(H)$ in (a), then LPO holds.

18. Let C be a convex subset of $\mathbb{C}^n$ such that $\sup\{f(z) : z \in C\}$ exists for each real (bounded) linear functional f on $\mathbb{C}^n$. Prove that C is totally bounded. (*Hint:* Prove that C is bounded; that

$$d(x) = \sup\left\{\inf_{y\in C} \operatorname{Re} f(x-y) : f \in S^*\right\}$$

exists for each $x \in \mathbb{C}^n$, where S^* is the unit ball of the dual of $\mathbb{C}^n$; and that $\rho(x, C)$ exists and equals $d(x)$.)

19. Let X, Y be normed spaces. We say that ϕ is *ultraweakly continuous* if it is uniformly continuous on $\mathcal{B}_1(X, Y)$ relative to the locally convex topology τ_w defined on $\mathcal{B}(X, Y)$ by the seminorms $T \rightsquigarrow y^*(Tx)$ with $x \in X$ and y^* a normed linear functional on Y. Prove that the following conditions are equivalent:

(a) $\mathcal{B}_1(X, Y)$ is τ_w-totally bounded.

(b) Every ultraweakly continuous linear functional on $\mathcal{B}(X, Y)$ is normed.

(c) Every ultraweakly continuous linear functional on $\mathcal{B}(X, Y)$ is compact.

Notes

A classical method of establishing the continuity of the boundary crossing map in Proposition 5.1.5 is to use a contradiction argument to prove that the mapping is sequentially continuous on $-C$; see [87] (pages 271–272). Our argument, based on the simple Lemma 5.1.3, is much more transparent and shorter than either that classical argument or its constructive counterpart (which is given in [32] and proves only the sequential continuity, not the full continuity, of the boundary crossing map). More information about boundary crossing can be found in [36], from which Exercises 2–4 are taken.

The Hahn–Banach theorem can be proved without recourse to the separation theorem (Exercise 6). Using the Church–Markov–Turing thesis, Metakides et al. [71] have produced an explicit example in which the hypotheses of the Hahn–Banach theorem hold but the linear functional cannot be extended with exact preservation of its norm. This makes Theorem 5.3.7 (which, together with Exercises 9 and 10, is due to Ishihara [55]) all the more significant, especially as it applies to many important examples of normed spaces, such as the L_p spaces for $p > 1$.

The locally convex topology on a finite-dimensional space X is unique and equivalent to the topology induced by any norm on X. The proof of this depends on some intricate geometric algorithms and is found in [39].

Finding necessary and sufficient conditions for the existence of the norm of a bounded linear mapping $T : X \longrightarrow Y$ between normed spaces, other than the case $Y = \mathbb{K}$, is an important unsolved problem, even in the case where T is an operator on a Hilbert space.

The Banach–Alaoglu theorem (Theorem 5.4.7) without the hypothesis of separability is classically equivalent to the *ultrafilter principle*—every filter is contained in an ultrafilter (see [80], page 766)—and is therefore unlikely to be provable constructively; see Exercise 1.4 of Chapter 1.

The introduction of the numbers r_{jk} in the proof of Proposition 5.4.16 is occasioned by the problem with the modulus–argument decomposition discussed at the start of Chapter 2.

Proposition 5.4.16 is a step in the direction of a constructive proof of the classical characterisation of weak-operator continuous linear functionals as those of the form $T \rightsquigarrow \sum_{k=1}^{n} \langle Tx_k, y_k \rangle$ with the vectors x_k, y_k in H. All attempts to generalise the technique used to prove Proposition 5.4.13 in order to characterise ultraweakly continuous linear functionals—those linear functionals on $\mathcal{B}(H)$ that are weak-operator uniformly continuous on $\mathcal{B}_1(H)$—suffer from a curse of dimensionality, failing because the n appearing in expressions like $\sum_{k=1}^{n} \langle Tx_k, y_k \rangle$ is usually more than 1. Although Exercise 16 requires only the weak-operator uniform continuity of the functional on the unit ball of $\mathcal{B}(H)$ and derives a stronger conclusion than that of Proposition 5.4.16, the solution that we have in mind for that exercise works only for a separable Hilbert space. In that case we can improve the conclusion to find square-summable sequences $(x_n)_{n \geqslant 1}$ and $(y_n)_{n \geqslant 1}$ in H such that for each $T \in \mathcal{B}(H)$ we have $\phi(T) = \sum_{n=1}^{\infty} \langle Tx_n, y_n \rangle$; see [23].

If H is a Hilbert space, then $\mathcal{B}_1(H)$ is classically weak-operator compact; however, the completeness of $\mathcal{B}_1(H)$ relative to the uniform structure associated with the weak-operator topology is an essentially nonconstructive property; see [14].

The result in Exercise 18 first appeared in [56].

6

Operators and Locatedness

Location! Location! Location!

—Unknown Estate Agent

We begin by introducing normed spaces on which the norm is differentiable in some fashion. In Section 2, with substantial help from the λ-technique, we provide criteria for the locatedness of certain convex sets in a normed space. This work is applied in Section 3 to proving that a bounded operator on a Hilbert space H has an adjoint if and only if it maps the unit ball to a located set. In the next section we construct approximate eigenvectors of a selfadjoint operator on H, and then show that a bounded positive operator has a unique positive square root. This result is applied in Section 5, in which we make further good use of the λ-technique to show that, for a so-called *weak-sequentially open* operator T on H, the range of T is located if and only if the range of its adjoint is located. The section ends with a proof of the closed range theorem for operators with an adjoint. The final section of the chapter presents a version of Baire's theorem, which is then applied to the proofs of three of the pillars of functional analysis: the open mapping, inverse mapping, and closed graph theorems.

6.1 Smooth and Uniformly Smooth Normed Spaces

As the reader will have observed, locatedness, classically a nugatory concept, plays an important role in many aspects of constructive analysis, such as the Hahn–Banach theorem and the separation theorem discussed in the preceding chapter. It is therefore fitting that criteria for locatedness should be a primary theme of this final chapter. Our first aim is to find necessary and sufficient conditions for the locatedness of convex sets in a normed space; this requires some preliminary results associated with the differentiability of the norm. Later in the chapter we shall apply our conditions to the locatedness of subsets associated with operators on a Hilbert space.

Throughout this section, unless we state otherwise, X will be a normed linear space. We say that X is *smooth* if its norm is Gâteaux differentiable at each nonzero vector; this is the case if and only if the limit

$$u_x(y) = \lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t} \tag{6.1}$$

exists for all unit vectors $x, y \in X$. We say X has *Fréchet differentiable norm* if this limit is uniform in the unit vector y: that is, if for each unit vector $x \in X$ and each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left| \frac{\|x + ty\| - \|x\|}{t} - u_x(y) \right| < \varepsilon \tag{6.2}$$

whenever $y \in X$, $\|y\| = 1$, and $0 < |t| < \delta$. If, moreover, δ can be chosen independent of x, then we say that X has *uniformly Fréchet differentiable norm* and is *uniformly smooth*. For this to be the case, for each $\varepsilon > 0$ there must exist $\delta > 0$ such that (6.2) holds whenever $x, y \in X$, $\|x\| = \|y\| = 1$, and $0 < |t| < \delta$.

With each vector x in a smooth normed space X we associate the mapping $J_x : X \longrightarrow \mathbb{R}$ defined by

$$J_x(y) = \lim_{t \to 0} \|x\| \frac{\|x + ty\| - \|x\|}{t}.$$

If $x = 0$, then $J_x(y) = 0$; if $x \neq 0$, then $J_x(y) = \|x\| \, u_x(y)$. In both these cases, $|J_x(y)| \leqslant \|x\| \, \|y\|$. In order to show that $J_x(y)$ is defined for an arbitrary element x of X, construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\begin{aligned} \lambda_n = 0 &\Longrightarrow \|x\| < \frac{1}{n}, \\ \lambda_n = 1 &\Longrightarrow \|x\| > \frac{1}{n+1}. \end{aligned}$$

We may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, set $\zeta_n = 0$; if $\lambda_n = 1$, set $\zeta_k = \|x\| \, u_x(y)$ for all $k \geqslant n$. Then $|\zeta_m - \zeta_n| \leqslant \|y\| / n$ whenever $m \geqslant n > 1$; so $(\zeta_n)_{n \geqslant 1}$ is a Cauchy sequence and therefore converges to a limit $\zeta \in \mathbb{R}$. Note that $|\zeta| \leqslant \|x\| \, \|y\|$. For each $\varepsilon > 0$, either $\|x\| \, \|y\| < \varepsilon / 2$ or $\|x\| \, \|y\| > 0$. In the first case we have

$$\left| \|x\| \frac{\|x + ty\| - \|x\|}{t} - \zeta \right| \leqslant 2 \|x\| \, \|y\| < \varepsilon$$

for all $t \neq 0$. In the second case, $\zeta = J_x(y)$ and

$$\left| \|x\| \frac{\|x + ty\| - \|x\|}{t} - \zeta \right| < \varepsilon$$

for all sufficiently small $t \neq 0$. Since $\varepsilon > 0$ is arbitrary, we conclude that $J_x(y)$ exists and equals ζ.

Proposition 6.1.1. *If X is smooth, then for each x in X, J_x is a normed linear functional on X, $\|J_x\| = \|x\|$, and $J_x(x) = \|x\|^2$. Moreover, if $c > 0$ and $x \in X$, then $J_{cx} = cJ_x$.*

Proof. If $x \neq 0$, then the first three conclusions follow from Proposition 5.3.6. On the other hand, if also $c > 0$, then

$$J_{cx} = \|cx\| \, u_{cx} = c \, \|x\| \, u_x = cJ_x.$$

The extension to the general case is left to the reader. □

Lemma 6.1.2. *If X is smooth, then for all $x, y \in X$,*

$$\|y\|^2 \leqslant \|x\|^2 - 2J_y \, (x - y) \, .$$

Proof. Using the linearity of J_y and also Proposition 6.1.1, we compute

$$J_y \, (x - y) = J_y(x) - \|y\|^2 \leqslant \|y\| \, \|x\| - \|y\|^2 \, .$$

Hence

$$\|x\|^2 - \|y\|^2 - 2J_y \, (x - y) \geqslant \|x\|^2 + \|y\|^2 - 2 \, \|x\| \, \|y\| = (\|x\| - \|y\|)^2 \geqslant 0.$$

□

Lemma 6.1.3. *Let X be a smooth normed space, and x a unit vector in X. Let δ, ε be positive numbers such that if $\|y\| = 1$ and $0 < |t| < \delta$, then*

$$\left| \frac{\|x + ty\| - \|x\|}{t} - J_x(y) \right| < \varepsilon.$$

Then for such t and each unit vector y,

$$\|x + ty\| + \|x - ty\| < 2 + 2t\varepsilon.$$

Proof. We may assume that $0 < t < \delta$. If $\|y\| = 1$, then

$$\frac{\|x + ty\| - \|x\|}{t} - J_x(y) < \varepsilon$$

and

$$\frac{\|x - ty\| - \|x\|}{-t} - J_x(y) > -\varepsilon.$$

Hence

$$\|x + ty\| < 1 + tJ_x(y) + t\varepsilon$$

and

$$\|x - ty\| < 1 - tJ_x(y) + t\varepsilon,$$

from which the result follows. □

We store the following result for use in Section 2.

Proposition 6.1.4. *Let X be a smooth, uniformly convex Banach space, and let f be a normed linear functional on X. Then there exists a unique vector $x \in X$ such that $f = J_x$.*

Proof. First assume that f is nonzero. By Proposition 2.3.7, there exists a unit vector $y \in X$ such that $f(y) = \|f\|$. Applying Proposition 5.3.6 to $\|f\|^{-1} f$, we see that $f = \|f\| \, u_y$. Now let $x = \|f\| \, y$. Then $\|x\| = \|f\|$ and (since $u_{cy} = u_y$ for all $c > 0$)

$$f = \|f\| \, u_y = \|x\| \, u_{\|f\| y} = \|x\| \, u_x = J_x.$$

It remains to deal with the uniqueness. Let $f = J_z$ for some $z \in X$. Then, by Proposition 6.1.1, $\|z\| = \|J_z\| = \|f\| > 0$, so

$$\|f\| \, u_y = f = \|f\| \, u_z = \|f\| \, u_{\|z\|^{-1} z}$$

and therefore $u_{\|z\|^{-1} z} = u_y$. Since $\|z\|^{-1} z$ is a unit vector, the uniqueness part of Proposition 5.3.6 shows that $\|z\|^{-1} z = y$; whence

$$z = \|z\| \, y = \|f\| \, y = x.$$

This completes the proof in the case where f is nonzero; the general case is left as an exercise. $\square$

The next lemma simplifies the proof of the theorem following it.

Lemma 6.1.5. *Let f be a mapping of $\mathbb{R} - \{0\}$ into $\mathbb{R}$ such that*

(a) *if $0 < t' \leqslant t$, then $f(-t) \leqslant f(-t') \leqslant f(t') \leqslant f(t)$, and*

(b) *for each $\varepsilon > 0$ there exists $t > 0$ such that $f(t) - f(-t) < \varepsilon$.*

Then $\lim_{t \to 0} f(t)$ exists.

Proof. Construct a sequence $(t_n)_{n \geqslant 1}$ of positive numbers decreasing strictly to 0 such that $f(t_n) - f(-t_n) < 1/n$ for each n. In view of (a), for $m \geqslant n$ we have

$$0 < f(t_m) - f(t_n) \leqslant f(t_m) - f(-t_m) < \frac{1}{m}.$$

Hence $(f(t_n))_{n \geqslant 1}$ is a Cauchy sequence in $\mathbb{R}$ and so converges to a limit $l \in \mathbb{R}$ such that $0 \leqslant f(t_m) - l \leqslant 1/m$ for each m. Given $\varepsilon > 0$, choose $t > 0$ as in (b) and then N such that $t_N < t$ and $0 \leqslant f(t_N) - l < \varepsilon$. By (a), if $0 < |t| < t_N$, then $f(-t_N) \leqslant f(t) \leqslant f(t_N)$ and therefore $|f(t) - l| \leqslant f(t_N) - l < \varepsilon$. $\square$

We have already defined "uniformly convex" for normed spaces, in Chapter 2. To introduce this notion for the dual of a normed space X, we concentrate on those elements of X^* that have a norm. We say that X^* is *uniformly convex* if for each $\varepsilon > 0$ there exists δ with $0 < \delta < 1$ such that for all normed $u, v \in X^*$ with $\|u\| = \|v\| = 1$, if there exists a unit vector $\xi \in X$ such that $\frac{1}{2}(u+v)(\xi) > 1 - \delta$, then $(u - v)(x) < \varepsilon$ for all unit vectors x.

Proposition 6.1.6. *A separable normed space is uniformly smooth if and only if it has a uniformly convex dual.*

Proof. Let X be a separable normed space, and suppose to begin with that X^* is uniformly convex. Given unit vectors $x, y \in X$, for all nonzero $t \in \mathbb{R}$ define

$$f(t) = \frac{\|x + ty\| - \|x\|}{t}.$$

Then f satisfies (a) of Lemma 6.1.5. It follows that in order to prove that X is uniformly smooth, we need only show that for each $\varepsilon > 0$ there exists $t > 0$, independent of the unit vectors x and y, such that $f(t) - f(-t) < \varepsilon$. To this end, let $\delta > 0$ be as in the definition of "uniformly convex" above, fix $t > 0$ such that

$$8t(2 + t) < \delta,$$

and consider any unit vectors $x, y \in X$. Note that $\|x + ty\| > 1/2$ and $\|x - ty\| > 1/2$. Defining unit vectors ξ, η by

$$\xi = \frac{1}{\|x + ty\|}(x + ty), \qquad \eta = \frac{1}{\|x - ty\|}(x - ty),$$

we obtain

$$\begin{aligned} \|\xi - \eta\| &= \frac{1}{\|x + ty\|\,\|x - ty\|}\,\|\|x - ty\|\,(x + ty) - \|x + ty\|\,(x - ty)\| \\ &\leqslant 4\,\|(\|x - ty\| - \|x + ty\|)\,x + t\,(\|x - ty\| + \|x + ty\|)\,y\| \\ &\leqslant 4\,[|\|x - ty\| - \|x + ty\|| + t\,(\|x - ty\| + \|x + ty\|)] \\ &\leqslant 4\,(2t + 2t(1 + t)) = 8t(2 + t) < \delta. \end{aligned}$$

Now let

$$\gamma = \min\left\{\frac{t\varepsilon}{4(1 + t)}, \frac{\delta}{2}\right\},$$

and use Proposition 5.3.1 to construct normed linear functionals u and v such that $\|u\| = \|v\| = 1$, $u(\xi) > 1 - \gamma$, and $v(\eta) > 1 - \gamma$. Then

$$(u + v)(\xi) = u(\xi) + v(\eta) + v(\xi - \eta) > 2 - 2\gamma - \|\xi - \eta\| \geqslant 2 - 2\frac{\delta}{2} - \delta = 2 - 2\delta$$

and therefore $\frac{1}{2}(u + v)(\xi) > 1 - \delta$. It follows that $(u - v)(z) < \varepsilon/2$ for all unit vectors $z \in X$, and in particular that $(u - v)(y) < \varepsilon/2$. Now,

$$u(x+ty) = \|x+ty\|\, u(\xi) > \|x+ty\|\,(1-\gamma)\,,$$

so

$$\|x+ty\| < u(x+ty) + \|x+ty\|\,\gamma \leqslant u(x+ty) + (1+t)\gamma \leqslant u(x+ty) + \frac{t\varepsilon}{4}.$$

Likewise,

$$\|x-ty\| < v(x-ty) + \frac{t\varepsilon}{4}.$$

Thus

$$\begin{aligned}
\frac{\|x+ty\|-\|x\|}{t} - \frac{\|x-ty\|-\|x\|}{-t} &= \frac{1}{t}\left(\|x+ty\| + \|x-ty\| - 2\right) \\
&\leqslant \frac{1}{t}\left(u(x+ty) + v(x-ty) + \frac{t\varepsilon}{2} - 2\right) \\
&= \frac{1}{t}\left(u(x) + v(x) + t(u-v)(y) + \frac{t\varepsilon}{2} - 2\right) \\
&\leqslant \frac{\varepsilon}{2} + (u-v)\,(y) < \varepsilon.
\end{aligned}$$

This completes the proof that X is uniformly smooth.

Now suppose, conversely, that X is uniformly smooth. Then for each $\varepsilon > 0$ there exists δ such that

$$0 < \delta < \min\left\{\frac{\varepsilon}{4}, 1\right\}$$

and such that for all unit vectors $x, y \in X$ and all t with $0 < |t| < 4\delta/\varepsilon$,

$$\left| \frac{\|x+ty\|-\|x\|}{t} - J_x(y) \right| < \frac{\varepsilon}{4}.$$

Letting $t = 2\delta/\varepsilon$, we see from Lemma 6.1.3 that for all unit vectors $x, y \in X$,

$$\|x+ty\| + \|x-ty\| < 2 + \frac{t\varepsilon}{2}.$$

Now consider normed $u, v \in X^*$ with $\|u\| = \|v\| = 1$, and assume that $\frac{1}{2}(u+v)(\xi) > 1-\delta$ for some unit vector $\xi \in X$. Then for all unit vectors $x \in X$ we have

$$\begin{aligned}
2 - \frac{t\varepsilon}{2} &= 2 - \delta < (u+v)(\xi) \\
&= u(\xi+tx) + v(\xi-tx) - t(u-v)(x) \\
&\leqslant \|\xi+tx\| + \|\xi-tx\| - t(u-v)(x) \\
&< 2 + \frac{t\varepsilon}{2} - t(u-v)(x)
\end{aligned}$$

and therefore $(u-v)(x) < \varepsilon$. Thus X^* is uniformly convex. □

If H is a Hilbert space, then, in view of the Riesz representation theorem, its dual is, like H itself, uniformly convex. It follows from Proposition 6.1.6 that a separable Hilbert space is uniformly smooth. In fact we can remove separability and completeness here.

Proposition 6.1.7. *An inner product space is uniformly smooth.*

Proof. Given unit vectors x, y in the inner product space X, for real t we have

$$\|x - ty\|^2 - \|x\|^2 = -2t \operatorname{Re} \langle x, y\rangle + t^2$$

and therefore

$$\|x - ty\| - \|x\| = \frac{-2t \operatorname{Re} \langle x, y\rangle + t^2}{\|x - ty\| + \|x\|}.$$

It follows that if $0 < |t| < 1/2$, then

$$\begin{aligned} \left| \frac{\|x - ty\| - \|x\|}{t} + \operatorname{Re} \langle x, y\rangle \right| &= \left| \frac{-2 \operatorname{Re} \langle x, y\rangle + t}{\|x - ty\| + \|x\|} + \operatorname{Re} \langle x, y\rangle \right| \\ &= \frac{1}{\|x - ty\| + 1} \left| (\|x - ty\| - 1) \operatorname{Re} \langle x, y\rangle + t \right| \\ &\leqslant \frac{2\,|t|}{3} (\operatorname{Re} \langle x, y\rangle + 1) \leqslant \frac{4\,|t|}{3}, \end{aligned}$$

which tends to 0 with t, independently of x and y. □

6.2 Locatedness of Convex Sets

Our aim in this section is to give conditions ensuring the locatedness of a bounded convex subset of a normed space under certain strong hypotheses that apply, in particular, to the norm on a Hilbert space or an L_p-space[1] for $1 < p < \infty$. Specifically, we aim to prove the following two results and some of their consequences.

Theorem 6.2.1. *Let X be a uniformly smooth normed space over $\mathbb{R}$, and C an inhabited, bounded, located convex subset of X. Then*

$$\sup \{J_x(y) : y \in C\}$$

exists for each $x \in X$.

[1] We refer the reader to [9, 12] for the theory of L_p spaces.

Theorem 6.2.2. *Let C be an inhabited, bounded, convex subset of a uniformly smooth normed space X over $\mathbb{R}$, such that*

$$\sup\{J_x(y) : y \in C\}$$

exists for each $x \in X$. Then C is located.

The path to Theorem 6.2.1 takes us through a tangle of technical lemmas, the first of which is an expression of the continuity of J_x as a function of x.

Lemma 6.2.3. *Let X be a real normed space with a Fréchet differentiable norm. For each unit vector $x \in X$ and each $\varepsilon > 0$ there exists $\delta > 0$ such that if $\|y\| < \delta$, then $|J_x(z) - J_{x+y}(z)| < \varepsilon$ for all unit vectors $z \in X$.*

Proof. Given a unit vector $x \in X$ and $\varepsilon > 0$, compute $\gamma \in (0,1)$ such that if $\|z\| = 1$ and $0 < |t| < \gamma$, then

$$\left| \frac{\|x + tz\| - \|x\|}{t} - J_x(z) \right| < \frac{\varepsilon}{8}. \tag{6.3}$$

Let

$$\delta = \min\left\{ \frac{\gamma\varepsilon}{32}, \frac{1}{2} \right\}$$

and consider a vector $y \in X$ with $\|y\| < \delta$. Noting that $\|x + y\| > 1/2$, set

$$\xi = \frac{1}{\|x + y\|}(x + y).$$

Then since

$$|1 - \|x + y\|| = |\|x\| - \|x + y\|| \leqslant \|y\|,$$

we have

$$\|x + y\| \|\xi - x\| = \|x + y - \|x + y\| x\| = \|(1 - \|x + y\|) x + y\| \leqslant 2\|y\|.$$

Also, by Proposition 6.1.1, $J_\xi(\xi) = \|\xi\|^2 = 1$, so

$$0 \leqslant 1 - J_\xi(x) = J_\xi(\xi - x) \leqslant \|\xi - x\| \leqslant \frac{2\|y\|}{\|x + y\|} < 4\delta. \tag{6.4}$$

On the other hand, setting $t = \gamma/2$ and applying Lemma 6.1.3, we see that for each unit vector $z \in X$,

$$\|x + tz\| + \|x - tz\| < 2 + \frac{t\varepsilon}{4}.$$

Therefore, by (6.4), the choice of t and δ, and Proposition 6.1.1,

$$\begin{aligned}
1-\frac{t\varepsilon}{4} &\leqslant 1-4\delta \\
&< J_\xi(x) \\
&= J_x(x)+J_\xi(x)-1 \\
&= J_x(x+tz)+J_\xi(x-tz)-1-t(J_x-J_\xi)(z) \\
&\leqslant \|x+tz\|+\|x-tz\|-1-t(J_x-J_\xi)(z) \\
&< \frac{t\varepsilon}{4}+1-t(J_x-J_\xi)(z).
\end{aligned}$$

Hence $(J_x - J_\xi)(z) < \varepsilon/2$. Replacing z by $-z$, we obtain $-(J_x - J_\xi)(z) < \varepsilon/2$ and therefore $|(J_x - J_\xi)(z)| < \varepsilon/2$. It follows from this and Proposition 6.1.1 that

$$\begin{aligned}
|(J_x-J_{x+y})(z)| &= |\|x+y\|\,(J_x-J_\xi)(z)+(\|x\|-\|x+y\|)\,J_x(z)| \\
&\leqslant \|x+y\|\,|(J_x-J_\xi)(z)|+|\|x\|-\|x+y\||\,|J_x(z)| \\
&\leqslant \frac{3}{2}\,|(J_x-J_\xi)(z)|+\|y\|\,\|x\|\,\|z\| \\
&< \frac{3\varepsilon}{4}+\frac{\varepsilon}{32}<\varepsilon,
\end{aligned}$$

as we wanted. □

Lemma 6.2.4. *Let X be a uniformly smooth normed space over $\mathbb{R}$, and let C be an inhabited, bounded, located convex subset of X. Let $x_0 \in X$ and $\varepsilon > 0$. Then there exists $\tau > 0$ such that for all $y \in C$,*

- ▷ *either there exists $x \in C$ such that $J_{x_0}(y) + \tau\varepsilon/6 < J_{x_0}(x)$*
- ▷ *or else $J_{x_0}(z) < J_{x_0}(y) + \varepsilon$ for all $z \in C$.*

Proof. Choose $M > 0$ such that $\|x - y\| < M$ for all $x, y \in C$. If $M\,\|x_0\| < \varepsilon$, then for all $x, y \in C$,

$$J_{x_0}(x-y) \leqslant \|x_0\|\,\|x-y\| \leqslant M\,\|x_0\| < \varepsilon.$$

Hence we may assume that $M\,\|x_0\| > 0$. From the definition of "uniformly smooth" and Lemma 6.2.3 we see that there exists $\delta > 0$ such that for all unit vectors $x, y \in X$,

- if $0 < |t| \leqslant \delta$, then

$$\left|\frac{\|x+ty\|-\|y\|}{t}-J_x(y)\right| < \frac{2\varepsilon}{3\,\|x_0\|\,M},$$

and

- if $\|y\| \leqslant \delta$, then for all unit vectors $z \in X$,

$$|J_x(z) - J_{x+y}(z)| < \frac{2\varepsilon}{3\|x_0\|M}.$$

Let

$$\tau = \min\left\{1, \frac{\delta\,\|x_0\|}{M}\right\}$$

and fix y in C. Either $\rho(x_0 + y, C) < \|x_0\|^2 - \tau\varepsilon/3$ or

$$\|x_0\|^2 - \frac{2\tau\varepsilon}{3} < \rho(x_0 + y, C). \tag{6.5}$$

In the first case, there exists $x \in C$ such that

$$\|x_0 + y - x\|^2 < \|x_0\|^2 - \frac{\tau\varepsilon}{3}.$$

It follows from Lemma 6.1.2 that

$$\|x_0\|^2 \leqslant \|x_0 + y - x\|^2 - 2J_{x_0}(y - x) < \|x_0\|^2 - \frac{\tau\varepsilon}{3} - 2J_{x_0}(y) + 2J_{x_0}(x)$$

and therefore

$$J_{x_0}(y) + \frac{\tau\varepsilon}{6} < J_{x_0}(x). \tag{6.6}$$

Thus it remains for us to show that if (6.5) holds, and therefore

$$\|x_0\|^2 - \frac{2\tau\varepsilon}{3} < \|x_0 + y - z\|^2 \text{ for all } z \in C, \tag{6.7}$$

then we can derive the second conclusion of our lemma. Consider any $x \in C$. If $\|x_0\|\,\|y - x\| < \varepsilon$, then

$$J_{x_0}(x) - J_{x_0}(y) = J_{x_0}(x - y) \leqslant \|x_0\|\,\|y - x\| < \varepsilon.$$

Thus without loss of generality we may assume that $\|x_0\|\,\|y - x\| > 0$. Since $(1 - \tau)\,y + \tau x$ belongs to the convex set C, we see from (6.7) and Proposition 6.1.1 that

$$\begin{aligned}
-\frac{2\tau\varepsilon}{3} &< \|x_0 + \tau\,(y - x)\|^2 - \|x_0\|^2 \\
&= J_{x_0+\tau(y-x)}\,(x_0 + \tau(y - x)) - \|x_0\|^2 \\
&= J_{x_0+\tau(y-x)}(x_0) + \tau J_{x_0+\tau(y-x)}(y - x) - \|x_0\|^2 \\
&\leqslant \|x_0 + \tau(y - x)\|\,\|x_0\| - \|x_0\|^2 + \tau J_{x_0+\tau(y-x)}(y - x).
\end{aligned}$$

Hence

$$0 < \|x_0\|\,\frac{\|x_0 + \tau(y - x)\| - \|x_0\|}{\tau} + J_{x_0+\tau(y-x)}(y - x) + \frac{2\varepsilon}{3}. \tag{6.8}$$

The appearance here of the expression

$$E = \frac{\|x_0 + \tau(y-x)\| - \|x_0\|}{\tau}$$

in the presence of a uniformly Fréchet differentiable norm suggests that we consider the unit vectors

$$u = \frac{1}{\|x_0\|} x_0, \quad v = \frac{1}{\|y-x\|}(y-x).$$

Writing

$$t = \frac{\tau \|y-x\|}{\|x_0\|},$$

we have

$$\frac{1}{\|y-x\|} E = \frac{\|u+tv\| - \|u\|}{t}.$$

Since $0 < t \leqslant \delta$, it follows that

$$\begin{aligned}
\frac{2\varepsilon}{3\|x_0\| M} &> \left| \frac{1}{\|y-x\|} E - J_u(v) \right| \\
&= \left| \frac{1}{\|y-x\|} E - \frac{1}{\|x_0\| \|y-x\|} J_{x_0}(y-x) \right| \\
&= \frac{1}{\|x_0\| \|y-x\|} \left| \|x_0\| E - J_{x_0}(y-x) \right|
\end{aligned}$$

and therefore

$$\left| \|x_0\| \frac{\|x_0 + \tau(y-x)\| - \|x_0\|}{\tau} - J_{x_0}(y-x) \right| < \frac{2\varepsilon \|y-x\|}{3M} < \frac{2\varepsilon}{3}. \tag{6.9}$$

In view of (6.8) and (6.9), it now makes sense to examine

$$\left| J_{x_0}(y-x) - J_{x_0+\tau(y-x)}(y-x) \right|,$$

which, by Proposition 6.1.1, equals

$$\|x_0\| \|y-x\| \left| J_u(v) - J_{u+tv}(v) \right|.$$

By our choice of δ, this last expression is less than

$$\|x_0\| \|y-x\| \frac{2\varepsilon}{3\|x_0\| M},$$

so

$$\left| J_{x_0}(y-x) - J_{x_0+\tau(y-x)}(y-x) \right| < \frac{2\varepsilon}{3}. \tag{6.10}$$

It follows from (6.8)–(6.10) that

$$\begin{aligned}0 &< \|x_0\| \frac{\|x_0+\tau(y-x)\|-\|x_0\|}{\tau} + J_{x_0+\tau(y-x)}(y-x) + \frac{2\varepsilon}{3} \\ &\leqslant \left(J_{x_0}(y-x)+\frac{2\varepsilon}{3}\right) + \left(J_{x_0}(y-x)+\frac{2\varepsilon}{3}\right) + \frac{2\varepsilon}{3},\end{aligned}$$

from which we obtain the inequality $J_{x_0}(x) < J_{x_0}(y) + \varepsilon$. Since x was an arbitrary point of C, the proof is complete. □

Using Lemma 6.2.4 and the λ-technique, we can now prove Theorem 6.2.1.

Proof. Under the hypotheses of Theorem 6.2.1, fix $M > 0$ such that $\|x-y\| \leqslant M$ for all $x, y \in C$, and consider any element x of X. It is enough to prove that for each $\varepsilon > 0$ there exists $y \in C$ such that

$$J_x(z) < J_x(y) + \varepsilon \qquad (z \in C). \tag{6.11}$$

For in that case, if $0 < \alpha < \beta$ and we choose $y \in C$ such that (6.11) holds with $\varepsilon = (\beta - \alpha)/2$, then either $J_x(y) > \alpha$, or else $J_x(y) < \frac{1}{2}(\alpha + \beta)$ and therefore $J_x(z) < \beta$ for all $z \in C$; so the desired supremum exists by the least-upper-bound principle.

Fixing $x \in X$, $y_0 \in C$, and $\varepsilon > 0$, let τ be as in Lemma 6.2.4. We construct an increasing binary sequence $(\lambda_n)_{n\geqslant 0}$ with $\lambda_0 = 0$, and a sequence $(y_n)_{n\geqslant 0}$ in C, such that

$$\begin{aligned}\lambda_n = 0 &\Longrightarrow J_x(y_{n-1}) + \frac{\tau\varepsilon}{6} < J_x(y_n), \\ \lambda_n = 1 &\Longrightarrow J_x(z) < J_x(y_{n-1}) + \varepsilon \text{ for all } z \in C.\end{aligned}$$

To do so, we assume that we have constructed $\lambda_0, \ldots, \lambda_n$ and $y_0, \ldots, y_n$. If $\lambda_n = 1$, set $\lambda_{n+1} = 1$ and $y_{n+1} = y_n$. If $\lambda_n = 0$, then by Lemma 6.2.4, either $J_x(y_n) + \tau\varepsilon/6 < J_x(y')$ for some $y' \in C$, or else $J_x(z) < J_x(y_n) + \varepsilon$ for all $z \in C$. In the first case set $\lambda_{n+1} = 0$ and $y_{n+1} = y'$. In the second case set $\lambda_{n+1} = 1$ and $y_{n+1} = y_n$. This completes the inductive construction.

Now choose N such that

$$M\|x\| < N\frac{\tau\varepsilon}{6}.$$

If $\lambda_N = 0$, then

$$J_x(y_0) + N\frac{\tau\varepsilon}{6} < J_x(y_N)$$

and therefore

$$N\frac{\tau\varepsilon}{6} < J_x(y_N - y_0) \leqslant M\|x\|,$$

a contradiction. Hence, in fact, $\lambda_N = 1$. □

Corollary 6.2.5. *Let X be a uniformly convex, uniformly smooth Banach space over $\mathbb{R}$, and let C be an inhabited, bounded, located convex subset of X. Then*

$$\sup\{f(y) : y \in C\}$$

exists for each normed linear functional f on X.

Proof. For a nonzero normed linear functional the result follows from Proposition 6.1.4 and Theorem 6.2.1. The completion of the proof is left to the reader. □

Turning now towards Theorem 6.2.2, we prove the following lemma.

Lemma 6.2.6. *Let C be an inhabited, bounded, convex subset of a uniformly smooth normed space X over $\mathbb{R}$ such that*

$$\sup\{J_x(z) : z \in C\}$$

exists for each $x \in X$, and let $\varepsilon > 0$. Then for each $x \in X$ there exists σ with $0 < \sigma < 1$ such that for each $y \in C$,

▷ *either there exists $y' \in C$ with $\|x - y'\| < (1 - \sigma)\|x - y\|$*

▷ *or else $\|x - y\| < \|x - z\| + \varepsilon$ for all $z \in C$.*

Proof. We may assume that $\varepsilon < 1$. Fixing $x \in X$, choose a positive integer $M > \max\{\varepsilon/4, 1\}$ such that $\|x - z\| < M$ for each $z \in C$. Since X is uniformly smooth, there exists δ with $0 < \delta < 2$ such that for all unit vectors $u, v \in X$,

$$0 < |t| < \delta \Longrightarrow \left| \frac{\|u - tv\| - \|u\|}{-t} - J_u(v) \right| < \frac{\varepsilon}{8M}. \tag{6.12}$$

Define

$$\sigma = \frac{\varepsilon\delta}{8M}$$

and note that $0 < \sigma < 1$. Let $y \in C$. If $\|x - y\| < \varepsilon$, then $\|x - y\| < \|x - z\| + \varepsilon$ for all $z \in C$. Hence we may assume that $\varepsilon/2 < \|x - y\|$. Now,

$$\sup\{J_{x-y}(z - y) : z \in C\} = \sup\{J_{x-y}(z) : z \in C\} - J_{x-y}(y)$$

exists, by our hypotheses. So either $J_{x-y}(z - y) < \varepsilon\|x - y\|$ for all $z \in C$ or else there exists $z_0 \in C$ such that

$$\frac{\varepsilon}{2}\|x - y\| < J_{x-y}(z_0 - y). \tag{6.13}$$

In the first case, for all $z \in C$ we have

$$\begin{aligned}\|x-y\|^2 &= J_{x-y}(x-y)\\ &= J_{x-y}(x-z)+J_{x-y}(z-y)\\ &< \|x-y\|\,\|x-z\|+\varepsilon\,\|x-y\|\,,\end{aligned}$$

and therefore

$$\|x-y\|<\|x-z\|+\varepsilon.$$

In the second case, setting

$$\tau=\frac{\delta\,\|x-y\|}{2M},\qquad y'=y+\tau(z_0-y),$$

we have $0<\tau<\delta/2<1$ and therefore $y'\in C$. Define also

$$u=\frac{1}{\|x-y\|}(x-y),\quad v=\frac{1}{\|z_0-y\|}(z_0-y),\quad t=\frac{\tau\|z_0-y\|}{\|x-y\|}.$$

Note, for the definition of v, that

$$\|z_0-y\|\,\|x-y\|\geqslant J_{x-y}(z_0-y)>\frac{\varepsilon}{2}\,\|x-y\|$$

and hence that $\|z_0-y\|>0$. Let

$$t=\frac{\tau\,\|z_0-y\|}{\|x-y\|}=\frac{\delta\,\|z_0-y\|}{2M}.$$

Then $t<\delta$ and, by Proposition 6.1.1 and (6.13),

$$tJ_u(v)=\frac{\tau}{\|x-y\|}J_u(z_0-y)=\frac{\tau}{\|x-y\|^2}J_{x-y}(z_0-y)>\frac{\tau\varepsilon}{2\,\|x-y\|}.$$

Hence

$$\begin{aligned}\|x-y'\| &= \|x-y-\tau(z_0-y)\|\\ &= \|x-y\|\,\|u-tv\|\\ &< \|x-y\|\left(\|u\|-tJ_u(v)+\frac{t\varepsilon}{8M}\right)\qquad\text{by (6.12)}\\ &< \|x-y\|-\frac{\tau\varepsilon}{2}+\frac{\tau\varepsilon\,\|z_0-y\|}{8M}.\end{aligned}$$

Since

$$\|z_0-y\|\leqslant\|z_0-x\|+\|x-y\|<2M$$

and (6.13) holds, this gives

$$\begin{aligned}\|x-y'\| &< \|x-y\|-\frac{\tau\varepsilon}{2}+\frac{\tau\varepsilon}{4}\\ &= \|x-y\|-\frac{\delta\varepsilon}{8M}\,\|x-y\|\\ &= (1-\sigma)\,\|x-y\|\end{aligned}$$

and completes the proof that, in this case, the first conclusion of the lemma obtains. □

We now prove Theorem 6.2.2.

Proof. Under the hypotheses of Theorem 6.2.2, let $x \in X$ and $\varepsilon > 0$. It suffices to show that there exists $y \in C$ such that $\|x - y\| < \|x - z\| + \varepsilon$ for all $z \in C$. Compute $\sigma > 0$ as in Lemma 6.2.6, and let $y_0 \in C$. If $\|x - y_0\| < \varepsilon$, then $\|x - y_0\| < \|x - z\| + \varepsilon$ for all $z \in C$. So we may assume that $\|x - y_0\| > 0$. Now construct an increasing binary sequence $(\lambda_n)_{n \geqslant 0}$ with $\lambda_0 = 0$, and a sequence $(y_n)_{n \geqslant 0}$ of elements of C, such that

$$\begin{aligned}\lambda_n = 0 &\Longrightarrow \|x - y_n\| < (1 - \sigma) \|x - y_{n-1}\| ,\\ \lambda_n = 1 &\Longrightarrow \|x - y_n\| < \|x - z\| + \varepsilon \text{ for all } z \in C.\end{aligned}$$

Suppose we have constructed $\lambda_0, \ldots, \lambda_n$ and $y_0, \ldots, y_n$. If $\lambda_n = 1$, set $\lambda_{n+1} = 1$ and $y_{n+1} = y_n$. If $\lambda_n = 0$, then by Lemma 6.2.6, either $\|x - y'\| < (1 - \sigma) \|x - y_n\|$ for some $y' \in C$, or else $\|x - y_n\| < \|x - z\| + \varepsilon$ for all $z \in C$. In the first case set $\lambda_{n+1} = 0$ and $y_{n+1} = y'$; in the second case set $\lambda_{n+1} = 1$ and $y_{n+1} = y_n$. Since $0 < \sigma < 1$, we see that $(1 - \sigma)^n \longrightarrow 0$ as $n \longrightarrow \infty$, so there exists $N > 1$ such that

$$(1 - \sigma)^n < \frac{\varepsilon}{\|x - y_0\|} \qquad (n \geqslant N) .$$

Now check $\lambda_0, \ldots, \lambda_N$. If there exists $n \leqslant N$ such that $\lambda_n = 1$, then we are finished; if $\lambda_n = 0$ for all $n \leqslant N$, then

$$\|x - y_N\| < (1 - \sigma)^N \|x - y_0\| < \varepsilon \leqslant \|x - z\| + \varepsilon$$

for all $z \in C$, and again we are finished. □

Corollary 6.2.7. *Let C be an inhabited, bounded, convex subset of a uniformly smooth normed space X over $\mathbb{R}$. Then C is located if and only if* $\sup \{J_x(z) : z \in C\}$ *exists for each $x \in X$.*

Proof. Apply Theorems 6.2.1 and 6.2.2. □

Corollary 6.2.8. *Let X be a uniformly convex, uniformly smooth Banach space over $\mathbb{R}$, and C an inhabited, bounded, convex subset of X. Then C is located if and only if* $\sup \{f(y) : y \in C\}$ *exists for each normed element f of X^*.*

Proof. Apply Proposition 6.1.4 and Corollary 6.2.8. □

Corollary 6.2.9. *Let C be an inhabited, bounded, convex subset of a Hilbert space H. Then C is located if and only if*

$$\sup \{\mathrm{Re} \langle x, y\rangle : y \in C\} \tag{6.14}$$

exists for each $x \in H$.

Proof. Consider first the case $\mathbb{K} = \mathbb{R}$. Since H is uniformly convex (see Section 3 of Chapter 4) and, by Proposition 6.1.7, uniformly smooth, the desired conclusion follows from the Riesz representation theorem and Corollary 6.2.8. We now easily obtain the result in the case $\mathbb{K} = \mathbb{C}$, since C is located in H if and only if it is located in the underlying real Hilbert space. $\square$

When we work with a Hilbert space rather than a general normed space, we can remove the boundedness hypothesis from Theorem 6.2.2. For this we need some preliminaries, including an important principle of functional analysis.

Lemma 6.2.10. *Let C be a convex subset of an inner product space X such that $\sup \{\mathrm{Re} \langle x, z\rangle : z \in C\}$ exists for each $x \in X$. Let $x \in X$ and $y \in C$. Let n be a positive integer and let $0 < \varepsilon < 1$. Then at least one of the following alternatives holds:*

(a) *There exists $y_0 \in C$ such that $\|x - y_0\|^2 < \|x - y\|^2 - \varepsilon^4/16n$.*

(b) *There exists $z_0 \in C$ such that $n/4 < \|x - y\|^2 + \|x - z_0\|^2$.*

(c) *$\|x - y\| < \|x - z\| + \varepsilon$ for all $z \in C$.*

Proof. If $\|x - y\| < \varepsilon$, then (c) holds trivially; so we may assume that $\|x - y\| > \varepsilon/2$. Our hypotheses ensure that

$$M = \sup \{\mathrm{Re} \langle x - y, z - y\rangle : z \in C\}$$

exists. Either $M < \varepsilon^2/2$ or else $\varepsilon^2/4 < M$. In the first case, for all $z \in C$ we have

$$\begin{aligned}
\|x - y\|^2 &= \mathrm{Re} \langle x - y, x - y\rangle \\
&= \mathrm{Re} \langle x - y, x - z\rangle + \mathrm{Re} \langle x - y, z - y\rangle \\
&< \|x - y\| \|x - z\| + \frac{\varepsilon^2}{2}
\end{aligned}$$

and therefore

$$\|x - y\| < \|x - z\| + \frac{\varepsilon^2}{2 \|x - y\|} < \|x - z\| + \varepsilon.$$

Hence (c) holds. We may therefore assume that $\varepsilon^2/4 < M$. Choosing $z_0 \in C$ such that

$$\tau = \operatorname{Re}\langle x - y, z_0 - y\rangle > \frac{\varepsilon^2}{4},$$

we consider three exhaustive cases.

Case 1: $\|z_0 - y\|^2 < 3\tau/2$. We have

$$\begin{aligned}
\|x - z_0\|^2 &= \|x - y - (z_0 - y)\|^2 \\
&= \|x - y\|^2 - 2\tau + \|z_0 - y\|^2 \\
&< \|x - y\|^2 - \frac{1}{2}\tau \\
&< \|x - y\|^2 - \frac{\varepsilon^2}{8} \\
&< \|x - y\|^2 - \frac{\varepsilon^4}{16n},
\end{aligned}$$

since $0 < \varepsilon < 1$. Hence (a) holds with $y_0 = z_0$.

Case 2: $n/2 < \|z_0 - y\|^2$. Using the parallelogram identity, we have

$$\begin{aligned}
\frac{n}{2} &< \|(z_0 - x) + (x - y)\|^2 \\
&= 2\|x - z_0\|^2 + 2\|x - y\|^2 - \|z_0 + y - 2x\|^2 \\
&\leqslant 2\|x - z_0\|^2 + 2\|x - y\|^2,
\end{aligned}$$

so (b) holds.

Case 3: $\tau < \|z_0 - y\|^2$ and $\|z_0 - y\|^2 < n$. Then

$$y_0 = y + \frac{\tau}{\|z_0 - y\|^2}(z_0 - y)$$

belongs to C, and

$$\begin{aligned}
\|x - y_0\|^2 &= \|x - y\|^2 - \frac{2\tau \operatorname{Re}\langle x - y, z_0 - y\rangle}{\|z_0 - y\|^2} + \frac{\tau^2}{\|z_0 - y\|^2} \\
&= \|x - y\|^2 - \frac{\tau^2}{\|z_0 - y\|^2} \\
&< \|x - y\|^2 - \frac{\varepsilon^4}{16n}
\end{aligned}$$

and so (a) holds. This completes the proof. □

The classical uniform boundedness theorem says:

If $(T_i)_{i\in I}$ is a family of bounded linear mappings from a Banach space X into a normed space Y such that for each $x \in X$ the family $(\|T_i x\|)_{i\in I}$ is bounded, then there exists $M > 0$ such that $\|T_i\| \leqslant M$ for each $i \in I$.

The reader is invited to consider why this form of the theorem is unlikely to be provable constructively. The correct constructive approach is via the contrapositive; here is a pretty version due to Royden [78].

Theorem 6.2.11. ***(Royden's uniform boundedness theorem)*** *Let $(T_n)_{n\geqslant 1}$ be a sequence of normed linear mappings from a Banach space X into a normed space Y such that $\|T_n\| > n3^n$ for each n. Then there exists $x \in X$ such that $\|T_n x\| > n$ for each n.*

Proof. For each $n \geqslant 1$ choose $x_n \in X$ such that $\|x_n\| = 3^{-n}$ and $\|T_n x_n\| > \frac{3}{4}\|T_n\|\|x_n\|$. Then

$$x = 4\sum_{n=1}^{\infty} x_n$$

is well defined, since (by comparison with $\sum_{n=1}^{\infty} 3^{-n}$) the series on the right converges in the complete space X. For $n \geqslant 2$, letting

$$y_n = x_1 + \cdots + x_{n-1},$$

we have

$$\begin{aligned}\frac{3}{2}\|T_n\|\|x_n\| &< 2\|T_n x_n\| = \|T(x_n + y_n) + T(x_n - y_n)\| \\ &\leqslant \|T_n(x_n + y_n)\| + \|T(y_n - x_n)\|,\end{aligned}$$

so at least one of the last two terms is greater than $\frac{3}{4}\|T_n\|\|x_n\|$. Replacing x_n by $-x_n$, if necessary, we may assume that

$$\|T_n(x_n + y_n)\| > \frac{3}{4}\|T_n\|\|x_n\|.$$

Hence

$$\begin{aligned}\frac{1}{4}\|T_n x\| &= \left\|T_n\left(\sum_{k=1}^{n} x_k\right) + T_n\left(\sum_{k=n+1}^{\infty} x_k\right)\right\| \\ &\geqslant \left\|T_n\left(x_n + \sum_{k=1}^{n-1} x_k\right)\right\| - \|T_n\|\sum_{k=n+1}^{\infty}\|x_k\| \\ &> \frac{3}{4}\|T_n\|3^{-n} - \|T_n\|\sum_{k=n+1}^{\infty} 3^{-k} \\ &\geqslant \|T_n\|3^{-n}\left(\frac{3}{4} - \frac{1}{2}\right) = \frac{1}{4}\|T_n\|3^{-n}\end{aligned}$$

and therefore $\|T_n x\| > 3^{-n}\|T_n\| > n$. □

Theorem 6.2.12. ***(Uniform boundedness theorem)*** *Let $(T_n)_{n\geqslant 1}$ be a sequence of normed linear mappings from a Banach space X into a normed space Y such that $\|T_n\| \longrightarrow \infty$ as $n \longrightarrow \infty$. Then there exists $x \in X$ such that the sequence $(\|T_n x\|)_{n\geqslant 1}$ is unbounded.*

Proof. Compute a strictly increasing sequence $(n_k)_{k\geqslant 1}$ of positive integers such that $\|T_{n_k}\| > k3^k$ for each k, and then apply Theorem 6.2.11. □

It is possible to remove from these two uniform boundedness theorems the condition that each T_n be normed; see Exercise 12.

Theorem 6.2.13. *Let C be an inhabited convex subset of a Hilbert space H such that*

$$\sup\{\operatorname{Re}\langle x,y\rangle : y \in C\}$$

exists for each $x \in H$. Then C is located.

Proof. Fix $\xi \in H$ and $y_0 \in C$. In order to prove that $\rho(\xi, C)$ exists, it is enough to show that if $0 < \varepsilon < 1$, then there exists $\eta \in C$ such that $\|\xi - \eta\| < \|\xi - y\| + \varepsilon$ for all $y \in C$. If $\|\xi - y_0\| < \varepsilon$, then we may take $\eta = y_0$. Hence we may assume that $\xi \neq y_0$. Setting $\lambda_0 = 0$, for the purposes of this proof we extend the λ-technique by constructing a sequence $(\lambda_n)_{n\geqslant 0}$ with values in $\{-1, 0, 1\}$, and a sequence $(y_n)_{n\geqslant 0}$ in C, such that for each $n \geqslant 1$,

$$\|\xi - y_n\| \leqslant \|\xi - y_{n-1}\|$$

and the following hold:

$$\begin{aligned}
\lambda_n = -1 &\Longrightarrow \|\xi - y_n\|^2 < \|\xi - y_{n-1}\|^2 - \varepsilon^4/16n, \\
\lambda_n = 0 &\Longrightarrow n/4 < \|\xi - y_n\|^2 + \|\xi - z\|^2 \text{ for some } z \in C, \\
\lambda_n = 1 &\Longrightarrow \|\xi - y_n\| < \|\xi - z\| + \varepsilon \text{ for all } z \in C.
\end{aligned}$$

Assume that we have already constructed $\lambda_0, \ldots, \lambda_n$ and $y_0, \ldots, y_n$. If $\lambda_n = 1$, set $\lambda_{n+1} = 1$ and $y_{n+1} = y_n$. If $\lambda_n \neq 1$, then we apply Lemma 6.2.10. If there exists $y \in C$ such that $\|\xi - y\|^2 < \|\xi - y_n\|^2 - \varepsilon^4/16n$, we set $\lambda_{n+1} = -1$ and $y_{n+1} = y$; if $n/4 < \|\xi - y_n\|^2 + \|\xi - z\|^2$ for some $z \in C$, we set $\lambda_{n+1} = 0$ and $y_{n+1} = y_n$; if $\|\xi - y_n\| < \|\xi - z\| + \varepsilon$ for all $z \in C$, we set $\lambda_{n+1} = 1$ and $y_{n+1} = y_n$. This completes our inductive construction of the sequences $(\lambda_n)_{n\geqslant 0}$ and $(y_n)_{n\geqslant 0}$. It remains to compute n such that $\lambda_n = 1$.

Let $(N_k)_{k\geqslant 1}$ be a strictly increasing sequence of positive integers such that for each k,

$$N_k > 4\left(k + \|\xi - y_0\|^2\right) \text{ and } \|\xi - y_0\|^2 < \sum_{n=N_k+1}^{N_{k+1}} \frac{\varepsilon^4}{16n}$$

(recall that $\sum_{n=1}^{\infty} 1/n$ diverges to infinity). Observe that if $\lambda_n = -1$ whenever $N_k < n \leqslant N_{k+1}$, then

$$0 \leqslant \left\|\xi - y_{N_{k+1}}\right\|^2 < \|\xi - y_{N_k}\|^2 - \sum_{n=N_k+1}^{N_{k+1}} \frac{\varepsilon^4}{16n} \leqslant \|\xi - y_0\|^2 - \sum_{n=N_k+1}^{N_{k+1}} \frac{\varepsilon^4}{16n} < 0,$$

a contradiction. Thus for each k there exists n_k such that $N_k < n_k \leqslant N_{k+1}$ and either $\lambda_{n_k} = 0$ or $\lambda_{n_k} = 1$. In the first case we choose $z' \in C$ such that

$$\frac{n_k}{4} < \|\xi - y_{n_k}\|^2 + \|\xi - z'\|^2$$

and set $w_k = \xi - z'$. In the second case set

$$w_k = \frac{n_k}{\|\xi - y_0\|} (\xi - y_0).$$

In each case we have

$$\|w_k\|^2 \geqslant \frac{n_k}{4} - \|\xi - y_0\|^2 > \frac{N_k}{4} - \|\xi - y_0\|^2 > k,$$

so $\|w_k\| \longrightarrow \infty$ as $k \longrightarrow \infty$. We now apply Corollary 6.2.12 to the normed linear functionals $x \rightsquigarrow \langle x, w_k \rangle$ on H, to produce $\xi_0 \in H$ such that the sequence $\{|\langle \xi_0, w_k \rangle|\}_{k \geqslant 1}$ is unbounded.

By our hypotheses, there exists $M > 0$ such that $|\langle \xi_0, \xi - z \rangle| < M$ for all $z \in C$. Choosing K such that $M < |\langle \xi_0, w_K \rangle|$, suppose that $\lambda_{n_K} = 0$. Then, by our construction of the vectors w_k, there exists $z \in C$ such that $w_K = \xi - z$ and therefore $|\langle \xi_0, w_K \rangle| = |\langle \xi_0, \xi - z \rangle| < M$, a contradiction. Hence $\lambda_{n_K} = 1$. □

We shall use some of these results on locatedness in the next section, when we look at adjoints of operators.

6.3 Adjoints

Throughout the rest of the chapter, H will be a Hilbert space. By an *operator* on H we mean a linear mapping of H into itself.

We observed in Section 3 of Chapter 5 that the constructive Riesz representation theorem, with its requirement that the linear functional be not just bounded but normed, does not enable us to prove that every bounded operator on H has an adjoint. In fact, we showed that the proposition "every bounded operator on a Hilbert space has an adjoint" implies LPO.

Are there general criteria for the existence of the adjoint? To answer this question affirmatively, we first prove a lemma.

Lemma 6.3.1. *Let T be an operator on H, and B a subset of H. Let a be a unit vector in H, and P the projection of H on the* 1*-dimensional subspace $\mathbb{K}a$. Then $PT(B)$ is located if and only if the set $\{\langle Tx, a\rangle : x \in B\}$ is located in $\mathbb{K}$.*

Proof. For all $x, y \in H$, since $(\mathrm{I} - P)\, y$ is orthogonal to $P(H)$, we have

$$\begin{aligned}\|y - PTx\|^2 &= \|Py - PTx\|^2 + \|(\mathrm{I} - P)\, y\|^2 \\ &= \|\langle Py, a\rangle\, a - \langle PTx, a\rangle\, a\|^2 + \|(\mathrm{I} - P)\, y\|^2 \\ &= |\langle y, a\rangle - \langle Tx, a\rangle|^2 + \|(\mathrm{I} - P)\, y\|^2 .\end{aligned}$$

Hence

$$\inf_{x \in B} \|y - PTx\|^2$$

exists if and only if

$$\inf_{x \in B} |\langle y, a\rangle - \langle Tx, a\rangle|^2$$

exists, from which the desired conclusion follows. □

Proposition 6.3.2. *Let T be a jointed operator on H. Let B be the unit ball of H, and P the projection of H on a* 1*-dimensional subspace. Then $PT(B)$ is located.*

Proof. Choose a unit vector $a \in H$ such that $Px = \langle x, a\rangle\, a$ for all $x \in H$. By the preceding lemma, it suffices to prove that the set

$$S = \{\langle Tx, a\rangle : x \in B\} = \{\langle x, T^*a\rangle : x \in B\}$$

is located in $\mathbb{K}$. We do this by showing that S is dense in the (located) ball $\overline{B}_{\mathbb{K}}(0, \|T^*a\|)$. Given $\varepsilon > 0$, we have either $\|T^*a\| < \varepsilon$, in which case $|\zeta - 0| < \varepsilon$ for all $\zeta \in \mathbb{K}$ with $|\zeta| \leqslant \|T^*a\|$, or else $T^*a \neq 0$. In the latter case, for each $\zeta \in \mathbb{K}$ with $\zeta \leqslant \|T^*a\|$ we have

$$\zeta = \left\langle \frac{\zeta}{\|T^*a\|^2} T^*a, T^*a \right\rangle$$

and

$$\frac{\zeta}{\|T^*a\|^2} T^*a \in B.$$

Thus in either case there exists $x \in B$ such that $|\zeta - \langle x, T^*a\rangle| < \varepsilon$. Since $\varepsilon > 0$ is arbitrary, it follows that S is dense in $\overline{B}_{\mathbb{K}}(0, \|T^*a\|)$. □

Proposition 6.3.3. *Let H be a Hilbert space with unit ball B, and T a bounded operator on H such that $PT(B)$ is located for each 1-dimensional projection P. Then T has an adjoint.*

Proof. By Lemma 6.3.1, for each unit vector $y \in H$, the set

$$C = \{\langle Tx, y\rangle : x \in B\}$$

is located in $\mathbb{K}$. Since T is bounded, C is a subset of a ball in $\mathbb{K}$. It follows from Corollary 4.1.7 and Proposition 2.2.10 that C is totally bounded. Thus, by Propositions 2.2.6 and 2.2.5, the linear functional u defined on H by

$$u(x) = \langle Tx, y\rangle \quad (x \in H)$$

is normed. By the Riesz representation theorem, there exists a unique element T^*y of H such that $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for all $x \in H$. □

Theorem 6.3.4. *Let T be a jointed operator on a Hilbert space H. Then the image under T of the unit ball is located.*

Proof. By Proposition 6.3.2 and Lemma 6.3.1, for each $x \in H$ the set

$$\{\langle Ty, x\rangle : y \in B\}$$

is located in $\mathbb{K}$. This set is also bounded, since

$$|\langle Ty, x\rangle| = |\langle y, T^*x\rangle| \leqslant \|T^*x\|$$

for all $y \in B$. Hence, as in the previous proof, it is totally bounded. Since the map $\zeta \rightsquigarrow \operatorname{Re}\zeta$ is uniformly continuous on $\mathbb{K}$, we now see that

$$\sup\{\operatorname{Re}\langle x, Ty\rangle : y \in B\} = \sup\{\operatorname{Re}\langle Ty, x\rangle : y \in B\}$$

exists for each $x \in H$. It follows from Theorem 6.2.13 applied to $C = T(B)$ that $T(B)$ is located in H. □

Theorem 6.3.5. *Let T be a bounded operator on a Hilbert space H that maps the unit ball to a located set. Then T has an adjoint.*

Proof. Given $y \in H$, we take $C = T(B)$ in Corollary 6.2.9 to show that the linear functional $x \rightsquigarrow \langle Ty, x\rangle$ is normed. The result now follows from the Riesz representation theorem, as in the proof of Proposition 6.3.3. □

Thus for a *bounded* operator on a Hilbert space, the existence of the adjoint is equivalent to the image of the unit ball being located.

When is a jointed operator bounded? The classical answer is "always"; the constructive answer is less decisive.

Theorem 6.3.6. ***(Hellinger–Toeplitz theorem)*** *Every jointed operator on a Hilbert space is sequentially continuous.*

Proof. Let T be a jointed operator on a Hilbert space H. It will suffice to prove that T^* is sequentially continuous, since we can then interchange T and T^* to obtain the desired result. Accordingly, let $(x_n)_{n\geqslant 1}$ be a sequence converging to 0 in H, and let $\varepsilon > 0$. By Ishihara's second trick (Lemma 3.2.2), either $\|T^*x_n\| < \varepsilon$ eventually or else there exists a strictly increasing sequence $(n_k)_{k\geqslant 1}$ of positive integers such that $\|T^*x_{n_k}\| > \varepsilon/2$ for all k. In the latter case, passing to a subsequence if necessary, we may assume that

$$\|x_{n_k}\| < \frac{\varepsilon}{2k^2 3^k}$$

for each k. Then, setting

$$y_k = \frac{2k3^k}{\varepsilon} x_{n_k},$$

we have $\|y_k\| < 1/k$ and $\|T^*y_k\| > k3^k$. We now apply Royden's uniform boundedness theorem (Theorem 6.2.11) to the normed linear functionals $x \rightsquigarrow \langle x, T^*y_k\rangle$ on H, to construct a vector $x \in H$ such that $|\langle x, T^*y_k\rangle| > k$ for each k. Since

$$\langle x, T^*y_k\rangle = \langle Tx, y_k\rangle \leqslant \|Tx\| \, \|y_k\| \longrightarrow 0 \text{ as } k \longrightarrow \infty,$$

this is absurd. We conclude that $\|T^*x_n\| < \varepsilon$ for all sufficiently large n. □

A subset S of $\mathbb{N}$ is *pseudobounded* if $\lim_{n\to\infty} n^{-1}s_n = 0$ for each sequence $(s_n)_{n\geqslant 1}$ in S. The following principle is trivially true in classical mathematics, holds in both INT and RUSS, and appears not to be provable in BISH (see [59]).

BD-$\mathbb{N}$ *Every inhabited, countable, pseudobounded set of positive integers is bounded.*

We now prove that if "sequentially continuous" can be replaced by "bounded" in the conclusion of Theorem 6.3.6, then **BD-**$\mathbb{N}$ holds.

Let

$$A = \{a_1, a_2, \ldots\}$$

be a countable pseudobounded set of positive integers. Let H be an infinite-dimensional Hilbert space with an orthonormal basis $(e_n)_{n\geqslant 1}$ of unit vectors. We first prove that for each $x \in H$,

$$Tx = \sum_{n=1}^{\infty} a_n \langle x, e_n\rangle \, e_n$$

is well defined. Let $(N_k)_{k\geqslant 1}$ be a strictly increasing sequence of positive integers such that

$$\sum_{n=N_k}^{\infty} |\langle x, e_n \rangle|^2 < 2^{-k-1}k^{-2},$$

and construct a binary sequence $(\lambda_k)_{k \geqslant 1}$ such that

$$\lambda_k = 0 \Longrightarrow \sum_{n=N_k}^{N_{k+1}-1} a_n^2 |\langle x, e_n \rangle|^2 < 2^{-k},$$

$$\lambda_k = 1 \Longrightarrow \sum_{n=N_k}^{N_{k+1}-1} a_n^2 |\langle x, e_n \rangle|^2 > 2^{-k-1}.$$

Define a sequence $(b_k)_{k \geqslant 1}$ in A as follows. If $\lambda_k = 0$, set $b_k = a_1$. If $\lambda_k = 1$, then

$$2^{-k-1}k^{-2} \max \left\{ a_n^2 : N_k \leqslant n < N_{k+1} \right\} \geqslant \sum_{n=N_k}^{N_{k+1}-1} a_n^2 |\langle x, e_n \rangle|^2 > 2^{-k-1},$$

so $a_n > k$ for some n with $N_k \leqslant n < N_{k+1}$; in this case we set $b_k = a_n$ for this n. Since A is pseudobounded, there exists m such that $b_k/k < 1$ for all $k \geqslant m$. If $\lambda_k = 1$ for some $k \geqslant m$, then $1 < b_k/k < 1$, a contradiction; hence $\lambda_k = 0$ for all $k \geqslant m$, and therefore the series defining Tx converges. It is easily seen that T is a one-one selfadjoint linear mapping of H onto itself; so, by Theorem 6.3.6, it is sequentially continuous. But if T is bounded, then the pseudobounded set A is bounded (by any positive bound for T).

6.4 Functions of Selfadjoint Operators

We say that a sequence $(S_n)_{n \geqslant 1}$ of operators on our Hilbert space H *converges strongly* to an operator S if

$$Sx = \lim_{n \to \infty} S_n x$$

for each $x \in H$. Given an operator T on H, for each power series

$$p(t) = \sum_{n=0}^{\infty} c_n t^n \tag{6.15}$$

with complex coefficients c_n we can form the corresponding power series in T,

$$p(T) = \sum_{n=0}^{\infty} c_n T^n,$$

which makes sense provided the series on the right converges strongly. In particular, if $p(t)$ is a polynomial, regarded as a power series whose coefficients are eventually 0, then $p(T)$ is always defined.

If each coefficient c_n in (6.15) is real and T is selfadjoint, then $p(T)$ is also selfadjoint. Our first aim is to prove that if, in that case, T has bound 1, then $p(T)$ is bounded by the sup norm of p on the interval $[0, 1]$. This will require us to prove some lemmas about approximate eigenvalues.

Note that two operators S, T on H are said to *commute* if $ST = TS$. Also, $T^0 = \mathrm{I}$ and $T^n = TT^{n-1}$ for any $n \geqslant 1$.

Lemma 6.4.1. *Let T be a selfadjoint operator on H with bound 1, and let x be a unit vector in H. Let*

$$v = Tx - \langle Tx, x\rangle\, x, \qquad u = \left\| x + \frac{1}{2}v \right\|^{-1} \left(x + \frac{1}{2}v \right).$$

Then

$$\langle Tu, u\rangle > \langle Tx, x\rangle + \frac{1}{4}\left\| v \right\|^2.$$

Moreover, if S is any operator that commutes with T, then $\|Su\| \leqslant 2\|Sx\|$.

Proof. Clearly, $v \in \{x\}^{\perp}$ and $Tx - v \perp \{x\}^{\perp}$, so (by Theorem 4.3.1) v is the projection of Tx onto $\{x\}^{\perp}$. Hence

$$\left\| x + \frac{1}{2}v \right\|^2 = 1 + \frac{1}{4}\left\| v \right\|^2 > 0,$$

and our definition of u makes sense. Also,

$$\langle Tx, v\rangle = \langle Tx, v\rangle - \langle \langle Tx, x\rangle\, x, v\rangle = \|v\|^2$$

and therefore, since $T = T^*$,

$$\begin{aligned}
\left\langle T\left(x + \frac{1}{2}v\right), x + \frac{1}{2}v \right\rangle &= \langle Tx, x\rangle + \operatorname{Re}\langle Tx, v\rangle + \frac{1}{4}\langle Tv, v\rangle \\
&\geqslant \langle Tx, x\rangle + \|v\|^2 - \frac{1}{4}\|Tv\|\,\|v\| \\
&\geqslant \langle Tx, x\rangle + \frac{3}{4}\|v\|^2.
\end{aligned}$$

Hence

$$\begin{aligned}
\langle Tu, u\rangle - \langle Tx, x\rangle &= \left\| x + \frac{1}{2}v \right\|^{-2} \left\langle T\left(x + \frac{1}{2}v\right), x + \frac{1}{2}v \right\rangle - \langle Tx, x\rangle \\
&\geqslant \left(1 + \frac{1}{4}\|v\|^2\right)^{-1} \left(\langle Tx, x\rangle + \frac{3}{4}\|v\|^2\right) - \langle Tx, x\rangle \\
&= \left(1 + \frac{1}{4}\|v\|^2\right)^{-1} \left(\frac{3}{4}\|v\|^2 - \frac{1}{4}\|v\|^2 \langle Tx, x\rangle\right) \\
&\geqslant \left(1 + \frac{1}{4}\|v\|^2\right)^{-1} \frac{1}{2}\|v\|^2 \\
&\geqslant \frac{1}{4}\|v\|^2,
\end{aligned}$$

since $\|v\| \leqslant 2$.

Finally, let S commute with T. Then

$$\begin{aligned}
\|Su\| &\leqslant \left\| S\left(x + \frac{1}{2}v\right)\right\| \\
&= \left\| Sx - \frac{1}{2}\langle Tx, x\rangle Sx + \frac{1}{2}TSx \right\| \\
&\leqslant \left(1 + \frac{1}{2}\left|\langle Tx, x\rangle\right| + \frac{1}{2}\right)\|Sx\| \leqslant 2\|Sx\|.
\end{aligned}$$

□

We show how to construct approximate eigenvectors common to finitely many commuting selfadjoint operators. We deal first with the case of a single selfadjoint operator.

Lemma 6.4.2. *Let T be a selfadjoint operator on H with bound 1, and let x_1 be a unit vector in H. Let ε be a positive number, and N an integer greater than $32/\varepsilon^2$. Define sequences $(x_n)_{n\geqslant 1}, (y_n)_{n\geqslant 1}$ in H recursively by*

$$y_n = Tx_n - \langle Tx_n, x_n\rangle x_n, \quad x_{n+1} = \left\| x_n + \frac{1}{2}y_n \right\|^{-1}\left(x_n + \frac{1}{2}y_n\right).$$

Then $\langle Tx_{n+1}, x_{n+1}\rangle \geqslant \langle Tx_n, x_n\rangle$ for each n, and there exists $n \leqslant N$ such that $\|y_n\| < \varepsilon$.

Proof. By Lemma 6.4.1,

$$\langle Tx_{n+1}, x_{n+1}\rangle - \langle Tx_n, x_n\rangle \geqslant \frac{1}{4}\|y_n\|^2.$$

On the other hand, either $\|y_n\| < \varepsilon$ for some $n \leqslant N$, or else $\|y_n\| > \varepsilon/2$ for all $n \leqslant N$. In the latter case, Lemma 6.4.1 shows that

$$\langle Tx_{n+1}, x_{n+1}\rangle - \langle Tx_n, x_n\rangle \geqslant \frac{\varepsilon^2}{16}$$

for each $n \leqslant N$, and therefore that

$$\langle Tx_{N+1}, x_{N+1}\rangle \geqslant \langle Tx_1, x_1\rangle + \frac{N\varepsilon^2}{16} > -1 + 2 = 1,$$

which is absurd since x_{N+1} is a unit vector and T has bound 1. □

Lemma 6.4.3. *Let $T_1, \ldots, T_n$ be commuting selfadjoint operators on H with common bound* 1, *and let x be a unit vector in H. Then for each $\varepsilon > 0$ there exists a unit vector u such that $\langle T_1 u, u\rangle \geqslant \langle T_1 x, x\rangle - \varepsilon$ and*

$$\|T_k u - \langle T_k u, u\rangle\, u\| < \varepsilon \quad (1 \leqslant k \leqslant n)\,.$$

Proof. Noting that the case $n = 1$ has been disposed of in Lemma 6.4.2, let $n > 1$ and suppose that we have proved the desired result for $n-1$ commuting selfadjoint operators. Fix $\varepsilon > 0$, and choose a positive integer $N > 32/\varepsilon^2$. By our induction hypothesis, there exists a unit vector x_1 such that

$$\langle T_1 x_1, x_1\rangle \geqslant \langle T_1 x, x\rangle - \frac{\varepsilon}{2}$$

and

$$\|T_k x_1 - \langle T_k x_1, x_1\rangle\, x_1\| < \frac{\varepsilon}{2^N} \quad (1 \leqslant k \leqslant n-1)\,.$$

Taking $T = T_n$, define sequences $(x_k)_{k\geqslant 1}$ and $(y_k)_{k\geqslant 1}$ as in Lemma 6.4.2. By that lemma, there exists j $(1 \leqslant j \leqslant N)$ such that $\|T_n x_j - \langle T_n x_j, x_j\rangle\, x_j\| < \varepsilon$. Setting $u = x_j$, we see that if $j = 1$, then we are finished; so we may assume that $j > 1$. Then for $1 \leqslant k \leqslant n-1$, since the selfadjoint operator $T_k - \langle T_k x_1, x_1\rangle\, \mathrm{I}$ commutes with T_n, we see from the final part of Lemma 6.4.1 that

$$\begin{aligned}
\|T_k u - \langle T_k x_1, x_1\rangle\, u\| &\leqslant 2\, \|T_k x_{j-1} - \langle T_k x_1, x_1\rangle\, x_{j-1}\| \\
&\leqslant \cdots \\
&\leqslant 2^{j-1}\, \|T_k x_1 - \langle T_k x_1, x_1\rangle\, x_1\| \\
&< 2^{N-1} \frac{\varepsilon}{2^N} = \frac{\varepsilon}{2}.
\end{aligned} \tag{6.16}$$

On the other hand, noting that for all $z \in H$ and all $a, b \in \mathbb{R}$,

$$\|z - au\|^2 - \|z - bu\|^2 = (a-b)\,(a + b - 2\,\mathrm{Re}\,\langle z, u\rangle)\,,$$

we have

$$\begin{aligned}
\|T_k u - \langle T_k u, u\rangle\, u\|^2 &- \|T_k u - \langle T_k x_1, x_1\rangle\, u\|^2 \\
&= (\langle T_k u, u\rangle - \langle T_k x_1, x_1\rangle)\,(\langle T_k u, u\rangle + \langle T_k x_1, x_1\rangle - 2\,\langle T_k u, u\rangle) \\
&= -\,(\langle T_k u, u\rangle - \langle T_k x_1, x_1\rangle)^2 \\
&\leqslant 0.
\end{aligned}$$

It follows from this, (6.16), and our choice of u that $\|T_k u - \langle T_k u, u\rangle\, u\| < \varepsilon$ for $1 \leqslant k \leqslant n$. Finally, again using (6.16), we see that

$$\begin{aligned}
\langle T_1 u, u\rangle &\geqslant \langle T_1 x_1, x_1\rangle - |\langle (T_1 - \langle T_1 x_1, x_1\rangle)\, u, u\rangle| \\
&\geqslant \langle T_1 x_1, x_1\rangle - \|T_1 u - \langle T_1 x_1, x_1\rangle\, u\| \\
&> \langle T_1 x, x\rangle - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} \\
&= \langle T_1 x, x\rangle - \varepsilon.
\end{aligned}$$

Our induction is now complete. □

Proposition 6.4.4. *Let T be a selfadjoint operator on H with bound* 1. *Then for each polynomial p with real coefficients, the operator $p(T)$ has bound*

$$M = \sup\{|p(t)| : 0 \leqslant t \leqslant 1\}.$$

Proof. Write

$$p(t) = \sum_{n=0}^{N} c_n t^n,$$

where each $c_n \in \mathbb{R}$. Let x be a unit vector in H, and let $\varepsilon > 0$. Since $p(T)$ has only real coefficients, it is selfadjoint. By Lemma 6.4.3, there exists a unit vector $u \in H$ such that $\langle p(T)u, u\rangle \geqslant \langle p(T)x, x\rangle - \varepsilon$ and

$$\|Tu - \langle Tu, u\rangle\, u\| < \left(1 + \sum_{n=1}^{N} n\,|c_n|\right)^{-1} \varepsilon.$$

Taking $t = \langle Tu, u\rangle$, for $n > 1$ we compute

$$\begin{aligned}
\|T^n u - t^n u\| &\leqslant \left\|T\left(T^{n-1}u - t^{n-1}u\right)\right\| + |t|^{n-1}\,\|Tu - tu\| \\
&\leqslant \left\|T^{n-1}u - t^{n-1}u\right\| + \|Tu - tu\| \\
&\leqslant \cdots \\
&\leqslant n\,\|Tu - tu\|.
\end{aligned}$$

Hence

$$\begin{aligned}
\|p(T)u - p(t)u\| &\leqslant \sum_{n=1}^{N} |c_n|\,\|T^n u - t^n u\| \\
&\leqslant \left(\sum_{n=1}^{N} n\,|c_n|\right) \|Tu - \langle Tu, u\rangle\, u\| < \varepsilon
\end{aligned}$$

and therefore

$$\begin{aligned}
\langle p(T)x, x\rangle &\leqslant \langle p(T)u, u\rangle + \varepsilon \\
&\leqslant |p(t)| + |\langle (p(T)u - p(t)u)\,, u\rangle| + \varepsilon \\
&\leqslant M + \|p(T)u - p(t)u\| < M + 2\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, it follows that $\langle p(T)x, x\rangle \leqslant M$. Replacing p by p^2, we now have

$$\begin{aligned}
\|p(T)x\|^2 &= \left\langle p(T)^2 x, x\right\rangle = \left\langle p^2(T)x, x\right\rangle \\
&\leqslant \sup\left\{|p(t)|^2 : 0 \leqslant t \leqslant 1\right\} = M^2
\end{aligned}$$

and therefore $\|p(T)x\| \leqslant M$. □

An operator T on H is said to be *positive* if $\langle Tx, x\rangle \geqslant 0$ for all $x \in H$. In that case, T is selfadjoint, since for all $x, y \in H$ we have both $\langle Tx, x\rangle = \langle x, Tx\rangle$ and the *polarisation identity*

$$\begin{aligned}4\langle Tx, y\rangle &= \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle \\ &\quad + \mathrm{i}\langle T(x+\mathrm{i}y), x+\mathrm{i}y\rangle - \mathrm{i}\langle T(x-\mathrm{i}y), x-\mathrm{i}y\rangle\end{aligned}$$

(which is actually valid for any operator T on H). We introduce a partial order on the set of selfadjoint operators by writing $S \leqslant T$, or equivalently $T \geqslant S$, if and only if $T - S$ is a positive operator.

For our first result about positive operators we use a weak constructive substitute for the classical result that a series of positive real numbers is convergent if its partial sums form a bounded sequence.

Lemma 6.4.5. *Let* $\sum_{n=1}^{\infty} a_n$ *be a series of nonnegative terms whose partial sums form a bounded sequence. Then for each* $\varepsilon > 0$ *and each positive integer* n *there exists* $k \geqslant n$ *such that* $a_k < \varepsilon$.

Proof. Let $b > 0$ be an upper bound for the partial sums of the series. Given $\varepsilon > 0$, choose a positive integer N such that $N\varepsilon/2 > b$. If $a_k > \varepsilon/2$ for $n < k \leqslant n + N$, then

$$\sum_{k=1}^{n+N} a_k \geqslant \sum_{k=n+1}^{n+N} a_k > \frac{N\varepsilon}{2} > b,$$

a contradiction. Hence $a_k < \varepsilon$ for some k with $n < k \leqslant n + N$. □

Proposition 6.4.6. *Let* S *and* T *be commuting positive operators on* H, *with* S *bounded. Then* ST *is positive.*

Proof. The proof is similar to the classical one on pages 415–417 of [4], so we give only an outline, leaving the details as an exercise. Since S is bounded, we may assume that $0 \leqslant S \leqslant \mathrm{I}$. Define a sequence $(S_n)_{n\geqslant 1}$ of bounded selfadjoint operators on H such that

$$S_1 = S, \qquad S_{n+1} = S_n - S_n^2 \quad (n \geqslant 1).$$

By induction, $0 \leqslant S_n \leqslant \mathrm{I}$ for each n. We now have

$$0 \leqslant \sum_{k=1}^{n} S_k^2 = \sum_{k=1}^{n} (S_k - S_{k+1}) = S - S_{n+1} \leqslant S.$$

It follows that for each $x \in H$,

$$\sum_{k=1}^{n} \|S_k x\|^2 = \sum_{k=1}^{n} \left\langle S_k^2 x, x\right\rangle = \langle (S - S_{n+1})x, x\rangle \leqslant \langle Sx, x\rangle .$$

By Lemma 6.4.5, there exists a strictly increasing sequence $(n_i)_{i\geqslant 1}$ of positive integers such that $\|S_{n_i}x\| < 2^{-i}$ for each i; whence

$$\sum_{k=1}^{n_i} S_k^2 x = Sx - S_{n_i+1}x \longrightarrow Sx \text{ as } i \longrightarrow \infty.$$

Since T commutes with S, it commutes with every S_n. Hence

$$\langle STx, x\rangle = \langle TSx, x\rangle = \left\langle T\left(\lim_{i\to\infty}\sum_{k=1}^{n_i} S_k^2 x\right), x\right\rangle.$$

Now, T is positive and hence selfadjoint; so, by Theorem 6.3.6, it is sequentially continuous. Hence

$$\begin{aligned}\langle STx, x\rangle &= \lim_{i\to\infty}\sum_{k=1}^{n_i}\left\langle TS_k^2 x, x\right\rangle \\ &= \lim_{i\to\infty}\sum_{k=1}^{n_i}\langle S_k T S_k x, x\rangle \\ &= \lim_{i\to\infty}\sum_{k=1}^{n_i}\langle TS_k x, S_k x\rangle \geqslant 0.\end{aligned}$$

Since x is arbitrary, we conclude that $ST \geqslant 0$. □

Our next objective is to construct the positive square root of a bounded positive operator. We first examine an iteration scheme for the function $t \rightsquigarrow \sqrt{t}$ on $[0,1]$. A standard classical proof of the convergence of that scheme uses Dini's highly nonconstructive theorem on the uniform convergence of monotone sequences of continuous functions [47] (pages 131–132). Fortunately, with relatively little extra effort, we can avoid using Dini's theorem altogether.

Lemma 6.4.7. *Define a sequence $(u_n)_{n\geqslant 1}$ of uniformly continuous mappings from $[0,1]$ into $\mathbb{R}$ iteratively by*

$$u_1(t) = 0, \quad u_{n+1}(t) = u_n(t) + \frac{1}{2}\left(t - u_n^2(t)\right).$$

Then $(u_n(t))_{n\geqslant 1}$ converges to $\sqrt{t}$ uniformly on $[0,1]$.

Proof. For each $t \in [0,1]$ and each n,

$$\begin{aligned}\sqrt{t} - u_{n+1}(t) &= \sqrt{t} - u_n(t) - \frac{1}{2}\left(t - u_n^2(t)\right) \\ &= \left(\sqrt{t} - u_n(t)\right)\left(1 - \frac{1}{2}\left(\sqrt{t} + u_n(t)\right)\right). \qquad (6.17)\end{aligned}$$

Using this, it is simple to prove by induction that $u_{n+1} \geqslant u_n$ and $u_n(t) \leqslant \sqrt{t}$. Let $0 < \varepsilon < 1$, and compute a positive integer N such that

$$\left(1 - \frac{\varepsilon}{2\sqrt{2}}\right)^n < \varepsilon \quad (n \geqslant N).$$

Consider any $t \in [0,1]$. If $t < \varepsilon^2$, then for all n we have $0 \leqslant u_n(t) \leqslant \sqrt{t} < \varepsilon$ and therefore

$$0 \leqslant \sqrt{t} - u_n(t) < \varepsilon. \tag{6.18}$$

If $t > \varepsilon^2/2$, then for all $n \geqslant 1$,

$$\frac{1}{2}\left(\sqrt{t} + u_n(t)\right) \geqslant \frac{1}{2}\sqrt{t} > \frac{\varepsilon}{2\sqrt{2}}$$

and therefore

$$1 - \frac{1}{2}\left(\sqrt{t} + u_n(t)\right) < 1 - \frac{\varepsilon}{2\sqrt{2}}.$$

It follows from this and (6.17) that for all $n \geqslant 2N$,

$$\begin{aligned}
0 \leqslant \sqrt{t} - u_n(t) &\leqslant \left(1 - \frac{\varepsilon}{2\sqrt{2}}\right)\left(\sqrt{t} - u_{n-1}(t)\right) \\
&\leqslant \left(1 - \frac{\varepsilon}{2\sqrt{2}}\right)^2\left(\sqrt{t} - u_{n-2}(t)\right) \\
&\leqslant \cdots \\
&\leqslant \left(1 - \frac{\varepsilon}{2\sqrt{2}}\right)^{n-N}\left(\sqrt{t} - u_N(t)\right) \\
&\leqslant \left(1 - \frac{\varepsilon}{2\sqrt{2}}\right)^N \\
&< \varepsilon.
\end{aligned}$$

Thus (6.18) holds for all $n \geqslant 2N$. □

Our next lemma will enable us to transform the iteration in Lemma 6.4.7 into one for the square root of a bounded positive operator.

Lemma 6.4.8. *Let T be a selfadjoint operator on H that satisfies $0 \leqslant T \leqslant \mathrm{I}$, define the sequence $(u_n)_{n\geqslant 1}$ as in* Lemma 6.4.7*, and write $U_n = u_n(T)$. Then for each n we have $0 \leqslant U_n \leqslant \frac{1}{2}(\mathrm{I} + T)$, $U_n^2 \leqslant T$, and*

$$T - U_{n+1}^2 = \left(T - U_n^2\right)\left(\mathrm{I} - \frac{1}{2}\left(U_{n+1} + U_n\right)\right). \tag{6.19}$$

Proof. We have

$$\begin{aligned} T - U_{n+1}^2 &= T - U_n^2 + \left(U_n^2 - U_{n+1}^2\right) \\ &= T - U_n^2 - \left(U_{n+1} - U_n\right)\left(U_{n+1} + U_n\right) \\ &= T - U_n^2 - \frac{1}{2}\left(T - U_n^2\right)\left(U_{n+1} + U_n\right) \\ &= \left(T - U_n^2\right)\left(\mathrm{I} - \frac{1}{2}\left(U_{n+1} + U_n\right)\right). \end{aligned}$$

Suppose that for some n we have $0 \leqslant U_n \leqslant \frac{1}{2}(\mathrm{I} + T)$ and $U_n^2 \leqslant T$. (These inequalities certainly hold for $n = 1$.) Then $0 \leqslant U_n \leqslant \mathrm{I}$, so U_n and $\mathrm{I} - U_n$ are positive and hence selfadjoint, and $(\mathrm{I} - U_n)^2 \geqslant 0$. Thus

$$\begin{aligned} U_{n+1} &= \frac{1}{2}(\mathrm{I} + T) - \frac{1}{2}\mathrm{I} + \left(U_n - \frac{1}{2}U_n^2\right) \\ &= \frac{1}{2}(\mathrm{I} + T) - \frac{1}{2}(\mathrm{I} - U_n)^2 \leqslant \frac{1}{2}(\mathrm{I} + T). \end{aligned}$$

On the other hand,

$$U_{n+1} - U_n = \frac{1}{2}\left(T - U_n^2\right) \geqslant 0$$

and therefore $U_{n+1} \geqslant U_n \geqslant 0$. Thus

$$0 \leqslant \frac{1}{2}(U_{n+1} + U_n) \leqslant U_{n+1} \leqslant \frac{1}{2}(\mathrm{I} + T) \leqslant \mathrm{I},$$

and so

$$\mathrm{I} - \frac{1}{2}(U_{n+1} + U_n) \geqslant 0.$$

Hence $T - U_{n+1}^2 \geqslant 0$, by (6.19) and Proposition 6.4.6. This completes the induction. □

Proposition 6.4.9. *Let T be a bounded positive operator on H. Then there exists a unique positive operator U on H such that $U^2 = T$. Moreover, U is bounded, U commutes with every operator that commutes with T, and the range of U is dense in the closure of the range of T.*

Proof. We may assume that $0 \leqslant T \leqslant \mathrm{I}$. Define $(u_n)_{n \geqslant 1}$ and $(U_n)_{n \geqslant 1}$ as in Lemmas 6.4.7 and 6.4.8. By Lemma 6.4.8, $0 \leqslant U_n \leqslant \mathrm{I}$. Since U_n is a polynomial in T, it commutes with every operator that commutes with T. By Lemma 6.4.7, for each $\varepsilon > 0$ there exists N_ε such that $|u_m(t) - u_n(t)| < \varepsilon$ for all $t \in [0,1]$ and all $m, n \geqslant N_\varepsilon$. It follows from Proposition 6.4.4 that $\|U_m x - U_n x\| \leqslant \varepsilon \|x\|$ for all $x \in H$ and all $m, n \geqslant N_\varepsilon$. Hence $(U_n)_{n \geqslant 1}$ converges strongly to an operator U on H. Clearly, $0 \leqslant U \leqslant \mathrm{I}$ and U commutes with every operator that commutes with T. Moreover, for each $n \geqslant 1$, since $u_n(t)$ is a strict polynomial (one without constant term) over $\mathbb{R}$, we have $\operatorname{ran}(U_n) \subset \operatorname{ran}(T)$, from which it follows that

$$\operatorname{ran}(U) \subset \overline{\operatorname{ran}(T)}. \tag{6.20}$$

With ε and N_ε as before, we see that

$$\begin{aligned}
\left|u_n^2(t) - t\right| &= \left|u_n(t) + \sqrt{t}\right| \left|u_n(t) - \sqrt{t}\right| \\
&\leqslant 2 \left|u_n(t) - \sqrt{t}\right| \\
&= 2 \lim_{m \to \infty} |u_n(t) - u_m(t)| \\
&\leqslant 2\varepsilon
\end{aligned}$$

for all $t \in [0,1]$ and all $n \geqslant N_\varepsilon$. It follows from Proposition 6.4.4 that

$$\left\| \left(U_n^2 - T\right) x \right\| \leqslant 2\varepsilon \, \|x\|$$

for all $x \in H$ and all $n \geqslant N_\varepsilon$. Hence

$$\begin{aligned}
\left\| \left(U^2 - T\right) x \right\| &\leqslant \left\| \left(U^2 - U_n^2\right) x \right\| + \left\| \left(U_n^2 - T\right) x \right\| \\
&= \left\| (U + U_n)(U - U_n) x \right\| + \left\| \left(U_n^2 - T\right) x \right\| \\
&\leqslant 2 \left\| (U - U_n) x \right\| + \left\| \left(U_n^2 - T\right) x \right\| \\
&\longrightarrow 0 \text{ as } n \longrightarrow \infty
\end{aligned}$$

and therefore $U^2 = T$. Moreover,

$$Tx = U \left(\lim_{n \to \infty} U_n x \right),$$

which, taken with (6.20), shows that $\operatorname{ran}(U)$ is dense in the closure of $\operatorname{ran}(T)$.

Now suppose that we have another positive operator S such that $S^2 = T$. Then

$$ST = SS^2 = S^2 S = TS,$$

so S commutes with T, and therefore U commutes with S. Given $x \in H$, write

$$y = Ux - Sx.$$

Then

$$\langle Uy, y \rangle + \langle Sy, y \rangle = \langle (U + S)(U - S) x, y \rangle = \left\langle \left(U^2 - S^2\right) x, y \right\rangle = 0.$$

But $\langle Uy, y \rangle \geqslant 0$ and $\langle Sy, y \rangle \geqslant 0$, so we must have $\langle Uy, y \rangle = 0 = \langle Sy, y \rangle$. Now apply the first part of the proof to U, to obtain a positive operator A such that $A^2 = U$. Then

$$\|Ay\|^2 = \langle Ay, Ay \rangle = \left\langle A^2 y, y \right\rangle = \langle Uy, y \rangle = 0,$$

so $Ay = 0$. Hence $Uy = A^2 y = 0$, and similarly, $Sy = 0$. It now follows that

$$\|Ux - Sx\|^2 = \langle (U - S)(U - S) x, x \rangle = \langle (U - S) y, x \rangle = 0.$$

Hence $S = U$. □

The operator U in Proposition 6.4.9 is called the *square root* of T and is denoted by $\sqrt{T}$ or $T^{1/2}$.

6.5 Locating the Kernel and the Range

Let T be a jointed operator on the Hilbert space H. In this section we deal with questions like the following:

> *When is* $\operatorname{ran}(T)$ *located? Is the locatedness of* $\operatorname{ran}(T)$ *linked to the locatedness of* $\ker(T)$ *or that of* $\operatorname{ran}(T^*)$? *When is* $\operatorname{ran}(T)$ *closed?*

As we shall see, the answers depend on some interesting applications of the λ-technique. We begin with two elementary lemmas.

Lemma 6.5.1. *Let X and Y be orthogonal subspaces of a Hilbert space H such that $X+Y$ is dense in H. Then both X and Y are located.*

Proof. Given $z \in H$, choose a sequence $(x_n)_{n\geqslant 1}$ in X and a sequence $(y_n)_{n\geqslant 1}$ in Y such that $x_n + y_n \longrightarrow z$ as $n \longrightarrow \infty$. For $m \geqslant n$ we have

$$\begin{aligned} \|x_m - x_n\|^2 + \|y_m - y_n\|^2 &= \|(x_m - x_n) + (y_m - y_n)\|^2 \\ &= \|(x_m + y_m) - (x_n + y_n)\|^2 \\ &\longrightarrow 0 \text{ as } n \longrightarrow \infty. \end{aligned}$$

Hence $(x_n)_{n\geqslant 1}, (y_n)_{n\geqslant 1}$ are Cauchy sequences in X, Y respectively, and so converge to respective limits x_∞, y_∞ in H such that $z = x_\infty + y_\infty$, $x_\infty \perp y_\infty$, and $y_\infty \perp X$. Given $x \in X$, we have

$$\|z - x\|^2 = \|y_\infty + (x_\infty - x)\|^2 = \|y_\infty\|^2 + \|x_\infty - x\|^2 \geqslant \|y_\infty\|^2,$$

with equality when $x = x_\infty$. It follows that $\rho(z, X)$ exists and equals $\|y_\infty\|$. Hence X, and likewise Y, is located. □

Lemma 6.5.2. *Let T be a jointed operator on H. Then* $\operatorname{ran}(T)^\perp = \ker(T^*)$. *Also,* $\operatorname{ran}(T)$ *is located if and only if* $\operatorname{ran}(T) + \ker(T^*)$ *is dense in H, in which case* $\ker(T^*)$ *is located.*

Proof. First observe that

$$y \perp \operatorname{ran}(T) \Longleftrightarrow \forall x \in H\, (\langle x, T^*y\rangle = 0) \Longleftrightarrow T^*y = 0,$$

so $\operatorname{ran}(T)^\perp = \ker(T^*)$. It follows that if $\operatorname{ran}(T)$ is located, then so is $\ker(T^*)$; moreover, if P is the projection of H on the closure of $\operatorname{ran}(T)$, then since $x = Px + (\mathrm{I} - P)x$ for each $\in H$, we see that $\operatorname{ran}(T) + \ker(T^*)$ is dense in H. If, conversely, $\operatorname{ran}(T) + \ker(T^*)$ is dense in H, then we can apply the preceding lemma to show that $\operatorname{ran}(T)$ and $\ker(T^*)$ are both located. □

Lemma 6.5.3. *If T is a jointed operator on H, then* $\operatorname{ran}(TT^*)$ *is dense in* $\operatorname{ran}(T)$.

Proof. Let B denote the closed unit ball of H. By Theorem 6.3.4, $TT^*(B)$ is located, so $TT^*(nB)$ is located for each positive integer n. Given $x \in H$ and $\varepsilon > 0$, we need only show that $\rho(Tx, TT^*(B_n)) < \varepsilon$ for some n. To this end, construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow \rho(Tx, TT^*(nB)) > \frac{\varepsilon}{2},$$
$$\lambda_n = 1 \Longrightarrow \rho(Tx, TT^*(nB)) < \varepsilon.$$

We may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, then, applying the separation and the Riesz representation theorems, we construct a unit vector $y_n \in H$ such that

$$\langle TT^*z, y_n \rangle + \frac{\varepsilon}{2} < \langle Tx, y_n \rangle \quad (z \in nB).$$

Then

$$n \left\| T^* y_n \right\|^2 < \langle TT^*(ny_n), y_n \rangle + \frac{\varepsilon}{2} < \langle Tx, y_n \rangle \leqslant \|Tx\|$$

and so $\left\| T^* y_n \right\|^2 < n^{-1} \|Tx\|$. If $\lambda_n = 1$, we set $y_n = 0$. Clearly, the sequence $(T^* y_n)_{n \geqslant 1}$ converges to 0. Choose a positive integer N such that $\langle x, T^* y_N \rangle < \varepsilon/2$. If $\lambda_N = 0$, then

$$\frac{\varepsilon}{2} \leqslant \frac{\varepsilon}{2} + \langle TT^*(Ny_N), y_N \rangle < \langle Tx, y_N \rangle = \langle x, T^* y_N \rangle < \frac{\varepsilon}{2},$$

a contradiction. Hence $\lambda_N = 1$. □

For a first application of Lemmas 6.5.2 and 6.5.3, we call an operator T on H *sequentially open* if for each sequence $(x_n)_{n \geqslant 1}$ such that $(Tx_n)_{n \geqslant 1}$ converges to 0, there exists a sequence $(y_n)_{n \geqslant 1}$ in $\ker(T)$ such that $x_n + y_n \longrightarrow 0$.

Proposition 6.5.4. *Let T be a sequentially open operator on H with an adjoint. Then* $\operatorname{ran}(T)$ *and* $\ker(T^*)$ *are located.*

Proof. By Lemma 6.5.3, there exists a sequence $(x_n)_{n \geqslant 1}$ in H such that

$$T(T^* x_n - x) \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Since T is sequentially open, there exists a sequence $(y_n)_{n \geqslant 1}$ in $\ker(T)$ such that $T^* x_n + y_n \longrightarrow x$ as $n \longrightarrow \infty$. We now see that $\operatorname{ran}(T^*) + \ker(T)$ is dense in H; whence, by Lemma 6.5.2, $\operatorname{ran}(T)$ is located. □

If T is any bounded operator with an adjoint, then T^*T is a bounded positive operator, so, by Proposition 6.4.9, the *absolute value* of T,

$$|T| = \sqrt{T^*T},$$

exists as a bounded positive operator.

Proposition 6.5.5. *If T is a bounded operator on H with an adjoint, then* $\operatorname{ran}(|T|)$ *is dense in the closure of* $\operatorname{ran}(T^*)$.

Proof. We see from Proposition 6.4.9 that the range of $\sqrt{T^*T}$ is dense in the closure of the range of T^*T. Reference to Lemma 6.5.3 completes the proof. □

We have already seen, in Lemma 6.5.2, that, for a jointed operator T, the kernel of T^* is the orthogonal complement of the range of T, and that if $\operatorname{ran}(T)$ is located, then so is $\ker(T^*)$. If, conversely, $\ker(T^*)$ is located, is $\operatorname{ran}(T)$ located also? Classically it is, since the closure of $\operatorname{ran}(T)$ is the orthogonal complement of $\ker(T^*)$. However, if the latter holds constructively, then we can prove Markov's principle: for if a is any real number such that $\neg(a=0)$, then the operator $T : z \rightsquigarrow az$ on $\mathbb{C}$ is selfadjoint and has kernel $\{0\}$; but if $\operatorname{ran}(T)$ is located, then $\neg(\rho(1,\operatorname{ran}(T))>0)$, so $\rho(1,\operatorname{ran}(T))=0$, $\operatorname{ran}(T)$ contains nonzero elements, and therefore $a \neq 0$.

We say that a sequence $(x_n)_{n\geqslant 1}$ in H *converges weakly* to $x \in H$, and we write

$$x_n \xrightarrow{w} x \quad \text{as } n \longrightarrow \infty$$

(or just $x_n \xrightarrow{w} x$), if $\lim_{n\to\infty} \langle x_n, y\rangle = \langle x, y\rangle$ for all $y \in H$. We call an operator T on H *weak-sequentially open* if for any sequence $(x_n)_{n\geqslant 1}$ such that $Tx_n \longrightarrow 0$, there exists a sequence $(y_n)_{n\geqslant 1}$ in $\ker(T)$ such that $x_n + y_n \xrightarrow{w} 0$.

Proposition 6.5.6. *If T is a jointed operator on H such that* $\operatorname{ran}(T^*)$ *is located, then T is weak-sequentially open.*

Proof. Let P be the projection of H on the closure of $\operatorname{ran}(T^*)$. Let $(x_n)_{n\geqslant 1}$ be a sequence in H such that $Tx_n \longrightarrow 0$, and for each n set $y_n = Px_n - x_n$. Then $y_n \in \operatorname{ran}(T^*)^\perp = \ker(T)$. For each $z \in H$ we have

$$\langle x_n + y_n, T^*z\rangle = \langle x_n, T^*z\rangle - \langle (\mathrm{I}-P)x_n, PT^*z\rangle = \langle Tx_n, z\rangle,$$

so

$$|\langle x_n + y_n, T^*z\rangle| \leqslant \|Tx_n\|\,\|z\| \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

It follows that for each $x \in H$ we have $\langle x_n + y_n, Px\rangle \longrightarrow 0$ and therefore

$$\langle x_n + y_n, x\rangle = \langle P(x_n+y_n), x\rangle = \langle x_n + y_n, Px\rangle \longrightarrow 0.$$

Hence T is weak-sequentially open. □

Theorem 6.5.7. *Let T be a weak-sequentially open, jointed operator on H with located kernel. Then T^* has located range.*

Proof. Let P be the projection of H on $\ker(T)$. It suffices to show that for each $x \in H$, the vector $x - Px$ is in the closure of $\operatorname{ran}(T^*)$: for then

$$\|x - y\|^2 = \|(x - Px) + y\|^2 + \|Px\|^2$$

for all $y \in \overline{\operatorname{ran}(T^*)}$, so $\rho\left(x, \operatorname{ran}(T^*)\right)$ exists and equals $\|Px\|$. Accordingly, fix x in H and $\varepsilon > 0$. Denote the closed unit ball in H by B. By Theorem 6.3.4, $T^*(nB)$ is located in H for each positive integer n; so we can construct an increasing binary sequence $(\lambda_n)_{n \geqslant 1}$ such that

$$\begin{aligned} \lambda_n = 0 &\Longrightarrow \rho\left(x - Px, T^*(nB)\right) > \frac{\varepsilon}{2}, \\ \lambda_n = 1 &\Longrightarrow \rho\left(x - Px, T^*(nB)\right) < \varepsilon. \end{aligned}$$

Without loss of generality we may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, then by Corollary 5.2.10 and the Riesz representation theorem, there exists a unit vector y_n such that for each $z \in nB$,

$$\langle x - Px, y_n \rangle > |\langle T^* z, y_n \rangle| + \frac{\varepsilon}{2} = |\langle z, T y_n \rangle| + \frac{\varepsilon}{2}.$$

Taking $z = n \left\|T y_n\right\|^{-1} T y_n$, we obtain

$$\frac{\varepsilon}{2} + n \left\|T y_n\right\| < \langle x - Px, y_n \rangle \leqslant \|x\|$$

and therefore $\left\|T y_n\right\| < n^{-1} \|x\|$. On the other hand, if $\lambda_n = 1 - \lambda_{n-1}$, we set $y_k = 0$ for all $k \geqslant n$. Clearly, the sequence $(T y_n)_{n \geqslant 1}$ converges to 0. But T is weak-sequentially open, so there exists a sequence $(z_n)_{n \geqslant 1}$ in $\ker(T)$ such that $y_n + z_n \xrightarrow{w} 0$. Choose N such that

$$|\langle x - Px, y_n \rangle| = |\langle x - Px, y_n + z_n \rangle| < \frac{\varepsilon}{2}$$

for all $n \geqslant N$. If $\lambda_N = 0$, then $|\langle x - Px, y_n \rangle| > \varepsilon/2$, which is absurd. Hence $\lambda_N = 1$. Since $\varepsilon > 0$ is arbitrary, we are through. □

Corollary 6.5.8. *If T is a jointed operator on H, then the following four statements are equivalent:*

(a) $\operatorname{ran}(T)$ *is located.*

(b) $\operatorname{ran}(T^*)$ *is located.*

(c) $\ker(T)$ *is located and T is weak-sequentially open.*

(d) $\ker(T^*)$ *is located and T^* is weak-sequentially open.*

Proof. This follows from Lemma 6.5.2, Theorem 6.5.7, and Proposition 6.5.6. □

Turning in a slightly different direction, we move towards a proof of the *closed range theorem*:

Theorem 6.5.9. *Let T be a jointed operator on H whose range is closed. Then* $\operatorname{ran}(T)$ *and* $\ker(T)$ *are located, and* $\operatorname{ran}(T^*)$ *is closed.*

To set the scene, consider how the closed range theorem is proved classically. One simple proof uses the polar decomposition of the operator T^*; but the existence of an exact polar decomposition in constructive analysis requires the range of the operator to be located from the outset (Exercise 16). In another classical proof of Theorem 6.5.9 the idea is to show that $\operatorname{ran}(T^*T)$ is closed and then to apply Lemma 6.5.3. To that end, let $(x_n)_{n\geqslant 1}$ be a sequence such that $(T^*Tx_n)_{n\geqslant 1}$ converges to a limit $y \in H$. Applying the classical uniform boundedness theorem (see page 163) to the bounded linear functionals f_n defined on the Hilbert space $\operatorname{ran}(T)$ by

$$f_n(Tx) = \langle Tx, Tx_n\rangle = \langle x, T^*Tx_n\rangle\,,$$

we obtain $M > 0$ such that $\|f_n\| \leqslant M$ for each n. Hence the linear functional $Tx \rightsquigarrow \langle x, y\rangle$ on $\operatorname{ran}(T)$ is bounded by M. By the Riesz representation theorem, there exists $x_\infty \in H$ such that

$$\langle x, y\rangle = \langle Tx, Tx_\infty\rangle = \langle x, T^*Tx_\infty\rangle$$

for all x. It follows that $y = T^*Tx_\infty \in \operatorname{ran}(T^*T)$.

This proof fails constructively in two places: first, in its use of the classical version of the uniform boundedness theorem, and second, in its application of the Riesz representation theorem, which requires the linear functional to be not just bounded but normed. Fortunately, as the following sequence of results will show, these difficulties can be overcome.

We begin with two lemmas that prepare us for a general result about sequentially continuous linear mappings between normed spaces.

Lemma 6.5.10. *Let $T : X \longrightarrow Y$ be a sequentially continuous linear mapping between normed spaces, $(x_n)_{n\geqslant 1}$ a Cauchy sequence in X, and $0 < \alpha < \beta$. Then either $\|Tx_n\| < \beta$ for all n or else there exists n such that $\|Tx_n\| > \alpha$.*

Proof. In view of the linearity of T, we may assume that $\beta - \alpha > 1$. Choosing a strictly increasing sequence $(N_k)_{k\geqslant 1}$ of positive integers such that $\|x_m - x_n\| < 2^{-3k}$ for all $m, n \geqslant N_k$, write

$$s_k = \max\{\|Tx_n\| : 1 \leqslant n \leqslant N_k\}\,.$$

Construct an increasing binary sequence $(\lambda_k)_{k\geqslant 1}$ such that

$$\lambda_k = 0 \Longrightarrow \forall j \leqslant k\ (s_j < \beta - 2^{-2j}),$$
$$\lambda_k = 1 \Longrightarrow \exists j \leqslant k\ (s_j > \beta - 2^{-2j+1}).$$

We may assume that $\lambda_1 = \lambda_2 = 0$. Now construct a sequence $(z_k)_{k\geqslant 1}$ in X as follows. If $\lambda_{k+1} = 0$ or $\lambda_k = 1$, set $z_k = 0$. If $\lambda_{k+1} = 1$ and $\lambda_k = 0$, then

$$\|Tx_{N_k}\| \leqslant s_k < \beta - 2^{-2k}$$

and $s_{k+1} > \beta - 2^{-2k-1}$, so we can choose j such that $N_k < j \leqslant N_{k+1}$ and $\|Tx_j\| > \beta - 2^{-2k-1}$; setting

$$z_k = 2^{2k}(x_j - x_{N_k}),$$

we have $\|z_k\| < 2^{-k}$. Moreover,

$$\begin{aligned}
\|Tz_k\| &= 2^{2k}\,\|Tx_j - Tx_{N_k}\| \\
&\geqslant 2^{2k}\left(\|Tx_j\| - \|Tx_{N_k}\|\right) \\
&> 2^{2k}\left(\beta - 2^{-2k-1} - (\beta - 2^{-2k})\right) = \frac{1}{2}.
\end{aligned}$$

This completes the construction of a sequence $(z_k)_{k\geqslant 1}$ converging to 0 in X. By the sequential continuity of T, $\lim_{k\to\infty} Tz_k = 0$. Choose K such that $\|Tz_k\| < 1/2$ for all $k \geqslant K$. If $\lambda_K = 1$, then there exists $n \leqslant N_K$ such that

$$\|Tx_n\| > \beta - 2^{-2n+1} > \alpha.$$

On the other hand, if $\lambda_K = 0$ and there exists $k \geqslant K$ such that $\lambda_{k+1} = 1 - \lambda_k$, then $\|Tz_k\| > 1/2$, a contradiction. Thus if $\lambda_K = 0$, then $\lambda_k = 0$ for all $k \geqslant K$ and therefore for all k, so $\|Tx_k\| < \beta$ for all k. □

Lemma 6.5.11. *Let $T : X \longrightarrow Y$ be a sequentially continuous linear mapping between normed spaces, and $(x_n)_{n\geqslant 1}$ a Cauchy sequence in X. Then* $\sup_{n\geqslant 1}\|Tx_n\|$ *exists.*

Proof. In view of the previous lemma, it is enough to show that the sequence $(\|Tx_n\|)_{n\geqslant 1}$ is bounded. To do so, choose $R > 0$ such that $\|x_n\| \leqslant R$ for all n. Taking $\alpha = 1$ and $\beta = 2$ in Lemma 6.5.10, we may assume that there exists n_1 such that $\|Tx_{n_1}\| > 1$. Set $\lambda_1 = 0$. Using Lemma 6.5.10 repeatedly, we construct an increasing binary sequence $(\lambda_n)_{n\geqslant 1}$, and an increasing sequence $(n_k)_{k\geqslant 1}$ of positive integers, such that

$$\lambda_k = 0 \Longrightarrow \|Tx_{n_k}\| > k \text{ and } n_k > n_{k-1},$$

$$\lambda_k = 1 \Longrightarrow (Tx_n)_{n\geqslant 1} \text{ is a bounded sequence and } n_{k+1} = n_k.$$

Assume that we have constructed λ_k and n_k. If $\lambda_k = 1$, we set $\lambda_{k+1} = \lambda_k$ and $n_{k+1} = n_k$. If $\lambda_k = 0$, then $\left\|Tx_{n_j}\right\| > j$ for all $j \leqslant k$. We then apply Lemma 6.5.10 to the Cauchy sequence $(x_j)_{j>n_k}$. Either we obtain $n_{k+1} > n_k$ such that $\left\|Tx_{n_{k+1}}\right\| > k+1$, or else $\|Tx_j\| < k+2$ for all $j > n_k$. In the first case we set $\lambda_{k+1} = 0$, and in the second, noting that $(Tx_n)_{n\geqslant 1}$ is bounded, we set $\lambda_{k+1} = 1$

and $n_{k+1} = n_k$. This completes the inductive construction of the sequences $(\lambda_k)_{k\geqslant 1}$ and $(n_k)_{k\geqslant 1}$. If $\lambda_k = 0$, set $z_k = k^{-1}x_{n_k}$; if $\lambda_k = 1$, set $z_k = 0$. Then $\|z_k\| \leqslant R/k$ for each k, so $z_k \longrightarrow 0$ and therefore, by the sequential continuity of T, $Tz_k \longrightarrow 0$. Choose K such that $\|Tz_k\| < 1$ for all $k \geqslant K$. If $\lambda_K = 0$, then

$$\|Tz_k\| = \frac{1}{k}\|Tx_{n_k}\| > 1,$$

a contradiction. Hence $\lambda_K = 1$ and so the sequence $(\|Tx_n\|)_{n\geqslant 1}$ is bounded. □

Proposition 6.5.12. *A sequentially continuous linear mapping $T : X \longrightarrow Y$ between normed spaces maps Cauchy sequences to Cauchy sequences.*

Proof. Given a Cauchy sequence $(x_n)_{n\geqslant 1}$ in X, choose a strictly increasing sequence $(N_k)_{k\geqslant 1}$ of positive integers such that $\|x_m - x_n\| < 2^{-k}$ for all $m, n \geqslant N_k$. For each k, the sequence $(x_n - x_{N_k})_{n\geqslant N_k}$ is a Cauchy sequence; so, by Lemma 6.5.11,

$$s_k = \sup_{n\geqslant N_k} \|Tx_n - Tx_{N_k}\|$$

exists. Given $\varepsilon > 0$, we construct an increasing binary sequence $(\lambda_k)_{k\geqslant 1}$ such that

$$\lambda_k = 0 \Longrightarrow s_k > \frac{\varepsilon}{4},$$
$$\lambda_k = 1 \Longrightarrow s_k < \frac{\varepsilon}{2}.$$

We may assume that $\lambda_1 = 0$. If $\lambda_k = 0$, choose $j \geqslant N_k$ such that $\|Tx_j - Tx_{N_k}\| > \varepsilon/4$, and set $z_k = x_j - x_{N_k}$. If $\lambda_k = 1$, set $z_k = 0$. Then $\|z_k\| < 2^{-k}$ for each k, so $z_k \longrightarrow 0$. Since T is sequentially continuous, $Tz_k \longrightarrow 0$ and we can choose K such that $\|Tz_k\| < \varepsilon/4$ for all $k \geqslant K$. If $\lambda_K = 0$, then $\|Tz_K\| > \varepsilon/4$, which is absurd; so $\lambda_K = 1$ and therefore $s_K < \varepsilon/2$. It follows that for all $j, k \geqslant N_K$,

$$\|Tx_j - Tx_k\| \leqslant \|Tx_j - Tx_{N_k}\| + \|Tx_k - Tx_{N_k}\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since ε is arbitrary, $(Tx_n)_{n\geqslant 1}$ is a Cauchy sequence in Y. □

We now move more directly towards the proof of the closed range theorem.

Lemma 6.5.13. *Let H be a Hilbert space, and T an operator on H with an adjoint and closed range. Let $(x_n)_{n\geqslant 1}$ be a sequence in H such that $(T^*Tx_n)_{n\geqslant 1}$ converges to* 0*. Then for all positive numbers α, β with $\alpha < \beta$, either $\|Tx_n\| > \alpha$ for some n or else $\|Tx_n\| < \beta$ for all n.*

Proof. Let $(n_k)_{k\geqslant 1}$ be a strictly increasing sequence of positive integers such that

$$\|T^*Tx_j\| < \frac{1}{(k+1)^2} \quad (j \geqslant n_k).$$

Taking $n_0 = 0$, for each $k \geqslant 1$ set

$$s_k = \max\{\|Tx_i\| : n_{k-1} < i \leqslant n_k\}.$$

Construct an increasing binary sequence $(\lambda_k)_{k\geqslant 1}$ such that

$$\lambda_k = 0 \Longrightarrow s_k < \beta,$$
$$\lambda_k = 1 \Longrightarrow s_k > \alpha.$$

We may assume that $\lambda_1 = 0$. Define a sequence $(y_k)_{k\geqslant 1}$ in $\operatorname{ran}(T)$ as follows: if $\lambda_k = 0$, set $y_k = 0$; if $\lambda_k = 1 - \lambda_{k-1}$, choose i with $n_{k-1} < i \leqslant n_k$ and $\|Tx_i\| > \alpha$, and set

$$y_j = \frac{1}{k\|Tx_i\|}Tx_i$$

for all $j \geqslant k$. Then $(y_k)_{k\geqslant 1}$ is a Cauchy sequence: in fact, $\|y_j - y_k\| \leqslant 1/(k+1)$ whenever $j \geqslant k$. Since $\operatorname{ran}(T)$ is closed in H and therefore complete, there exists $z \in H$ such that $(y_k)_{k\geqslant 1}$ converges to Tz. Choosing a positive integer N such that $\|z\| < N\alpha$, consider any integer $k \geqslant N$. If $\lambda_k = 1 - \lambda_{k-1}$, then

$$Tz = \frac{1}{k\|Tx_i\|}Tx_i$$

for some i with $n_{k-1} < i \leqslant n_k$ and $\|Tx_i\| > \alpha$, so

$$\frac{1}{k}\alpha < \frac{1}{k}\|Tx_i\| = \langle Tx_i, Tz\rangle = \langle T^*Tx_i, z\rangle \leqslant \frac{1}{k^2}\|z\| < \frac{1}{k^2}N\alpha \leqslant \frac{1}{k}\alpha,$$

a contradiction. Hence $\lambda_k = \lambda_{k-1}$ for all $k \geqslant N$. It follows that either $\lambda_k = 0$ for all k, or else $\lambda_k = 1 - \lambda_{k-1}$ for some $k \leqslant N$. In the first case, $\|Tx_n\| < \beta$ for all n; in the second, $\|Tx_i\| > \alpha$ for some i with $n_{k-1} < i \leqslant n_k$. □

Lemma 6.5.14. *Under the hypotheses of* Lemma 6.5.13*, for all positive numbers α, β with $\alpha < \beta$, either $\|Tx_n\| > \alpha$ for infinitely many n or else $\|Tx_n\| < \beta$ for all sufficiently large n.*

Proof. Let $(n_k)_{k\geqslant 1}$ be a strictly increasing sequence of positive integers such that

$$\|T^*Tx_j\| < \frac{1}{k^2} \quad (j \geqslant n_k).$$

Successively applying Lemma 6.5.13 to the sequences $(x_{n_k}, x_{n_k+1}, \ldots)$ for $k \geqslant 1$, construct an increasing binary sequence $(\lambda_k)_{k\geqslant 1}$ such that

$$\lambda_k = 0 \Longrightarrow \exists i \geqslant n_k (\alpha < \|Tx_i\|),$$
$$\lambda_k = 1 \Longrightarrow \forall i \geqslant n_k (\|Tx_i\| < \beta).$$

We may assume that $\lambda_1 = 0$. Define a sequence $(y_k)_{k\geqslant 1}$ in $\operatorname{ran}(T)$ as follows: if $\lambda_k = 0$, choose $i \geqslant n_k$ such that $\|Tx_i\| > \alpha$ and set

$$y_k = \frac{1}{k\|Tx_i\|} Tx_i;$$

if $\lambda_k = 1 - \lambda_{k-1}$, set $y_j = y_{k-1}$ for all $j \geqslant k$. Then $\|y_j - y_k\| \leqslant 2/k$ for all $j \geqslant k$; so $(y_k)_{k\geqslant 1}$ is a Cauchy sequence in the complete space $\operatorname{ran}(T)$ and therefore converges to Tz for some $z \in H$. Choosing a positive integer N such that $\|z\| < N\alpha$, consider any integer $k > N$. If $\lambda_k = 1 - \lambda_{k-1}$, then

$$Tz = \frac{1}{(k-1)\|Tx_i\|} Tx_i$$

for some $i \geqslant n_{k-1}$ with $\|Tx_i\| > \alpha$, so

$$\begin{aligned} \frac{1}{k-1}\alpha < \frac{1}{k-1}\|Tx_i\| &= \langle Tx_i, Tz\rangle = \langle T^*Tx_i, z\rangle \\ &\leqslant \frac{1}{(k-1)^2}\|z\| < \frac{1}{(k-1)^2} N\alpha \leqslant \frac{1}{k-1}\alpha, \end{aligned}$$

a contradiction. Hence $\lambda_k = \lambda_{k-1}$ for all $k > N$. It follows that either $\lambda_k = 1$ for some $k \leqslant N$ or else $\lambda_k = 0$ for all k. In the first case, $\|Tx_i\| < \beta$ for all $i \geqslant n_k$. In the second case, for each k there exists $i \geqslant n_k$ such that $\|Tx_i\| > \alpha$. □

Lemma 6.5.15. *Under the hypotheses of* Lemma 6.5.13, $(Tx_n)_{n\geqslant 1}$ *converges to* 0.

Proof. By Lemma 6.5.14, for each $\varepsilon > 0$ either $\|Tx_n\| > \varepsilon/2$ for infinitely many n or else $\|Tx_n\| < \varepsilon$ for all sufficiently large n. In the former case, passing to an appropriate subsequence, we may assume that $\|Tx_n\| > \varepsilon/2$ and $\|T^*Tx_n\| < 1/n^2$ for all n. Applying the uniform boundedness theorem (Corollary 6.2.12) to the normed linear functionals $Tx \rightsquigarrow \langle Tx, nTx_n\rangle$ on the Hilbert space $\operatorname{ran}(T)$, we can find $Tz \in \operatorname{ran}(T)$ such that the sequence $(|\langle Tz, nTx_n\rangle|)_{n\geqslant 1}$ is unbounded. But

$$|\langle Tz, nTx_n\rangle| = |\langle z, nT^*Tx_n\rangle| \leqslant \frac{1}{n}\|z\| \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

This contradiction rules out the possibility that $\|Tx_n\| > \varepsilon/2$ for infinitely many n. Since $\varepsilon > 0$ is arbitrary, the result follows. □

Lemma 6.5.16. *Let H be a Hilbert space, and T a jointed operator on H with closed range. Then* $\operatorname{ran}(T^*T)$ *is complete.*

Proof. First observe that, by Theorem 6.3.6, both T and T^* are sequentially continuous. It is easily seen using Lemma 6.5.2 that T^* is one-one on $\operatorname{ran}(T)$. So we can define a linear mapping

$$S : \operatorname{ran}(T^*T) \longrightarrow \operatorname{ran}(T)$$

by setting $ST^*Tx = Tx$ for each $x \in H$. Lemma 6.5.15 shows that S is sequentially continuous. Consider any Cauchy sequence $(T^*Tx_n)_{n\geqslant 1}$ in $\operatorname{ran}(T^*T)$. By Proposition 6.5.12, $(Tx_n)_{n\geqslant 1}$ is a Cauchy sequence in the Hilbert space $\operatorname{ran}(T)$. Hence there exists $x \in H$ such that $Tx_n \longrightarrow Tx$. The sequential continuity of T^* now yields $T^*Tx_n \longrightarrow T^*Tx$. Thus $\operatorname{ran}(T^*T)$ is complete. □

We can now give the proof of the closed range theorem.

Proof. By Lemmas 6.5.3 and 6.5.16, $\operatorname{ran}(T^*T)$ is both dense in $\operatorname{ran}(T^*)$ and complete. Hence $\operatorname{ran}(T^*)$ is complete and therefore closed in H. Moreover, for each $x \in H$ there exists $y \in H$ such that $T^*x = T^*Ty$; so

$$x = Ty + (x - Ty),$$

where $Ty \in \operatorname{ran}(T)$ and (by Lemma 6.5.2)

$$x - Ty \in \ker(T^*) = \operatorname{ran}(T)^{\perp}.$$

It follows from Lemma 6.5.2 that both $\operatorname{ran}(T)$ and $\ker(T^*)$ are located. Applying Corollary 6.5.8, we now see that $\operatorname{ran}(T^*)$ and $\ker(T)$ are located. □

6.6 Baire's Theorem, with Applications

Baire's theorem states that

> *The intersection of a sequence of dense open subsets of a complete metric space is dense in that space.*

The standard classical proof of this theorem (see [79], page 97) passes over unchanged to the constructive setting. However, various classically equivalent versions of Baire's theorem do not pass over unscathed; for example, the version that states that if a complete metric space is the union of a sequence of closed subsets, then one of those subsets is inhabited. In this section we present a constructive proof of a restricted form of this last version of Baire's theorem in the context of a Banach space, and apply it to operator theory.

The proof of our version of Baire's theorem introduces yet another technique, in which we show that a certain property P holds by constructing an element of the set $\{x \in X : P\}$, which is empty if $\neg P$ holds and inhabited if P holds.

Theorem 6.6.1. *Let X be a Banach space, and C a closed, convex, balanced, located subset of X such that $X = \bigcup_{n \geqslant 1} nC$ and the distance $\rho(0, -C)$ exists. Then C° is inhabited.*

Proof. For each positive integer n define the open set

$$U_n = -nC \cup \{x \in X : C^\circ \text{ is inhabited}\},$$

where $-nC$ is the metric complement of nC in X. To prove that U_n is dense in X, consider $y \in X$ and $\varepsilon > 0$. Note that $\rho(y, nC)$ exists and equals $n\rho\left(n^{-1}y, C\right)$. Either $\rho(y, nC) > 0$ or $\rho(y, nC) < \varepsilon$. In the first case, $y \in -nC$. In the second, choose $z \in nC$ such that $\|y - z\| < \varepsilon$. Noting that

$$\rho(0, -2nC) = 2n\rho(0, -C)$$

exists, we see that either $\rho(0, -2nC) < 2\varepsilon$ or $\rho(0, -2nC) > \varepsilon$. In the former case, choose $z' \in -2nC$ such that $\|z'\| < 2\varepsilon$. For each $w \in nC$ we have $-w \in nC$ (since C is balanced), so

$$z - w \in nC + nC = 2nC,$$

by the convexity of C. Hence

$$\|(z - z') - w\| = \|z' - (z - w)\| \geqslant \rho(z', 2nC) > 0.$$

Thus $z - z' \in -nC$. Since also

$$\|y - (z - z')\| \leqslant \|y - z\| + \|z'\| < 3\varepsilon,$$

we see that $\rho(y, U_n) < 3\varepsilon$. Finally, in the case $\rho(0, -2nC) > \varepsilon$, for each x with $\|x\| \leqslant \varepsilon$ we have $\rho(x, 2nC) = 0$. Hence

$$\overline{B}(0, \varepsilon) \subset \overline{2nC} = 2nC,$$

so $(2nC)^\circ$, and therefore C°, is inhabited. In this case, $U_n = X$. This completes the proof that U_n is dense in X.

Since X is complete, it follows from the standard version of Baire's theorem that $\bigcap_{n \geqslant 1} U_n$ is dense in X and therefore, in particular, contains a point ξ. Choose n such that $\xi \in nC$. Since also $\xi \in U_n$, we must have

$$\xi \in \{x \in X : C^\circ \text{ is inhabited}\}.$$

Hence C° is indeed inhabited. □

Recall the classical open mapping theorem for bounded linear mappings: a bounded linear mapping T of a Banach space X onto a Banach space Y is *open*, in the sense that it maps open subsets of X onto open subsets of Y (or, equivalently, it maps the open unit ball of X onto an open subset of Y). This theorem does not hold constructively without some additional hypotheses: the inverse S of the mapping T in the example on page 169 is a bounded operator of a Hilbert space onto a Hilbert space, but if S is open, then T is bounded. In other words, if every bounded linear mapping of a Hilbert space onto a Hilbert space is open, then **BD**-$\mathbb{N}$ holds.

Nonetheless, we can prove constructive versions of the open mapping theorem that are classically equivalent to the standard version. To that end, we begin with a lemma.

Lemma 6.6.2. *Let T be a sequentially continuous linear mapping of a Banach space X into a normed space Y such that $T(B_X(0,1))$ is located. Let r be a positive number, and y an element of $B_Y(0,r)$. There exists $x \in \overline{B}_X(0,2)$ such that if $y \neq Tx$, then $\rho(z, T(B_X(0,1))) > 0$ for some $z \in B_Y(0,r)$.*

Proof. If $\rho(y, T(B_X(0,1))) > 0$, then we can take x to be any element of $B_X(0,2)$. So we may assume that $\rho(y, T(B_X(0,1))) < r/2$. Choosing $x_1 \in B_X(0,2)$ such that $\left\| y - T\left(\frac{1}{2}x_1\right)\right\| < r/2$ and therefore $\|2y - Tx_1\| < r$, we set $\lambda_1 = 0$. This is the first step in the inductive construction of an increasing binary sequence $(\lambda_n)_{n\geqslant 0}$ and a sequence $(x_n)_{n\geqslant 0}$ of elements of $B_X(0,2)$ such that for each $n \geqslant 1$,

▷ if $\lambda_n = 0$, then

$$\rho\left(\left(2^{n-1}y - \sum_{i=1}^{n-1} 2^{n-1-i}Tx_i\right), T(B_X(0,1))\right) < \frac{r}{2}$$

and

$$\left\| 2^n y - \sum_{i=1}^{n} 2^{n-i}Tx_i \right\| < r;$$

▷ if $\lambda_n = 1 - \lambda_{n-1}$, then

$$\rho\left(\left(2^{n-1}y - \sum_{i=1}^{n-1} 2^{n-1-i}Tx_i\right), T(B_X(0,1))\right) > 0$$

and $x_i = 0$ for all $i \geqslant n$.

Suppose that we have found λ_{n-1} and x_{n-1} with the applicable properties. If $\lambda_{n-1} = 1$, we set $\lambda_n = 1$ and $x_n = 0$. If $\lambda_{n-1} = 0$, we consider the two cases

$$\rho\left(\left(2^{n-1}y - \sum_{i=1}^{n-1} 2^{n-1-i}Tx_i\right), T(B_X(0,1))\right) > 0$$

and

$$\rho\left(\left(2^{n-1}y-\sum_{i=1}^{n-1}2^{n-1-i}Tx_i\right),T\left(B_X\left(0,1\right)\right)\right)<\frac{r}{2}.$$

In the first case we set $\lambda_n = 1$ and $x_n = 0$. In the second case we choose $x_n \in B_X(0,2)$ such that

$$\left\|2^{n-1}y-\sum_{i=1}^{n-1}2^{n-1-i}Tx_i-T\left(\frac{1}{2}x_n\right)\right\|<\frac{r}{2}$$

and set $\lambda_n = 0$. Then

$$\left\|2^{n}y-\sum_{i=1}^{n}2^{n-i}Tx_i\right\|<r.$$

This completes the induction.

Since X is complete, the series $\sum_{i=1}^{\infty}2^{-i}x_i$ converges to a sum $x \in \overline{B}_X(0,2)$. Then $Tx=\sum_{i=1}^{\infty}2^{-i}Tx_i$, by the sequential continuity of T. If $y \neq Tx$, then there exists N such that

$$\left\|y-\sum_{i=1}^{N}2^{-i}Tx_i\right\|>2^{-N}r$$

and therefore

$$\left\|2^{N}y-\sum_{i=1}^{N}2^{N-i}Tx_i\right\|>r.$$

We must therefore have $\lambda_N = 1$; so there exists $n \leqslant N$ such that $\lambda_n = 1-\lambda_{n-1}$. Setting

$$z=2^{n-1}y-\sum_{i=1}^{n-1}2^{n-1-i}Tx_i,$$

we see that $\rho\left(z,T\left(B_X\left(0,1\right)\right)\right)>0$ (as $\lambda_n = 1$) and $\|z\|<r$ (as $\lambda_{n-1}=0$). This completes the proof. □

Lemma 6.6.3. *Let C be a balanced convex subset of a normed space Y, and let $y \in Y$ and $r > 0$ be such that $B(y,r) \subset C$. Then $B(0,r) \subset C$.*

Proof. If $z \in Y$ and $\|z\| < r$, then $y \pm z \in C$ and therefore $z=\frac{1}{2}(y+z)-\frac{1}{2}(y-z) \in C$. □

Here is our version of the *open mapping theorem.*

Theorem 6.6.4. *Let X, Y be Banach spaces, and T a sequentially continuous linear mapping of X onto Y such that $T(B(0,1))$ is located and $\rho(0,-T(B(0,1)))$ exists. Then T is an open mapping.*

Proof. Since

$$Y = \bigcup_{n \geqslant 1} n\left(\overline{T(B(0,1))}\right),$$

we can apply Theorem 6.6.1 to compute $y_0 \in Y$, $R > 0$, and a positive integer N such that

$$B_Y(y_0, R) \subset N\left(\overline{T(B(0,1))}\right).$$

Writing $y_1 = N^{-1}y_0$ and $r = R/N$, we obtain

$$B_Y(y_1, r) \subset \overline{T(B(0,1))}.$$

It follows from Lemma 6.6.3 that

$$B_Y(0, r) \subset \overline{T(B(0,1))}. \tag{6.21}$$

Now consider any $y \in Y$ with $\|y\| < r$. Choose $x \in \overline{B}(0,2)$ as in the conclusion of Lemma 6.6.2. If $y \neq Tx$, then there exists $z \in B_Y(0,r)$ such that $\rho(z, T(B(0,1))) > 0$, which contradicts (6.21); hence (the inequality on a normed space being tight) $y = Tx$. Thus $B_Y(0,r) \subset T(B(0,2))$, and therefore T is an open mapping. □

A subset C of a linear space X is said to be a *generating set* for, or to *generate,* X if every element of X is a finite linear combination of elements of C. Our next lemma will enable us to prove that compactly generated Banach spaces are finite-dimensional.

Lemma 6.6.5. *Let G be a compact generating set for a nontrivial Banach space X. Then there exists a compact generating set C for X that is both convex and balanced such that $\rho(0, -C)$ exists.*

Proof. We may assume that G is both convex and balanced (the proof is left as an exercise). Now, $2G$ is compact, the mapping $x \rightsquigarrow \rho(x, G)$ is uniformly continuous on X, and X is nontrivial. Hence there exists $\delta > 0$ such that both the sets

$$C = \{x \in 2G : \rho(x,G) \leqslant \delta\},$$
$$D = \{x \in 2G : \rho(x,G) \geqslant \delta\}$$

are compact. Note that C is convex and balanced, and, since it contains G, generates X. We show that $-C$ is dense in D. To this end, consider any $x \in D$ and any $\varepsilon > 0$. Choose $t > 1$ such that $(t-1)\|x\| < \varepsilon/2$, and suppose that $tx \in C$. Then, since C is balanced, $x \in C$; whence $x \in C \cap D$ and therefore $\rho(x, G) = \delta$. But then for each $g \in G$ we have

$$\|tx - g\| = t\left\|x - t^{-1}g\right\| \geqslant t\rho(x,G) = t\delta;$$

so $\rho(tx, G) \geqslant t\delta > \delta$, which is absurd since $tx \in C$. We conclude that $tx \notin C$. It follows from Proposition 3.1.2 that there exists $y \in {\sim}C$ such that $\|tx - y\| < \varepsilon/2$

and therefore $\|x-y\|<\varepsilon$. Applying Bishop's lemma (Proposition 3.1.1), we see that $y \in -C$. This completes the proof that $-C$ is dense in D. Since the norm function is uniformly continuous on the compact set D, it now follows that

$$\rho(0,-C)=\inf\{\|x\|: x \in -C\}=\inf\{\|x\|: x \in D\}$$

exists. □

A normed space with a compact generating set is said to be *compactly generated.*

Theorem 6.6.6. *A compactly generated Banach space is finite-dimensional.*

Proof. Let X be a compactly generated Banach space. We first suppose that X contains a nonzero vector. By Lemma 6.6.5, X has a balanced, convex, compact generating set C such that $\rho(0,-C)$ exists. Since $X=\bigcup_{n \geqslant 1} nC$, we can apply Theorem 6.6.1 to show that C° is inhabited; whence C contains a nontrivial ball. But every ball in a normed space is located, so the ball in question is totally bounded. It follows from Proposition 4.1.13 that X is finite-dimensional.

It remains to remove the restriction that X be nontrivial. To do this, we work in the product Banach space $X \times \mathbb{K}$. This space is generated by the compact set $G \times\{1\}$; so, by the foregoing, $X \times \mathbb{K}$ is finite-dimensional. It follows that X, being isomorphic to the quotient space $(X \times \mathbb{K}) / \mathbb{K}$, is finite-dimensional. □

Recall from page 102 that a linear mapping $T: X \longrightarrow Y$ between normed spaces is said to be *compact* if $T(B_X(0,1))$ is a totally bounded subset of Y.

Corollary 6.6.7. *If T is a compact linear mapping of a normed space X onto a Banach space Y, then Y is finite-dimensional.*

Proof. The totally bounded set $T(B(0,1))$ generates Y. Since Y is complete, it follows that $\overline{T(B(0,1))}$ is a compact generating set for Y; whence, by Theorem 6.6.6, Y is finite-dimensional. □

In the traditional development of functional analysis, the open mapping theorem is used to prove Banach's inverse mapping theorem, the closed graph theorem, and the uniform boundedness theorem. The last of these three we have already discussed, in Section 2. To deal with the inverse mapping theorem we need another lemma about convex sets, and two further technical lemmas.

Lemma 6.6.8. *Let C be a convex, absorbing subset of a Banach space X. Then $0 \notin \overline{-C}$.*

Proof. Assuming that $0 \in \overline{-C}$, we first show that $X - nC$ is dense in X for each positive integer n. To do so, construct a sequence $(x_n)_{n\geqslant 1}$ in $-C$ that converges to 0. Fixing a positive integer n, an element y of X, and $\varepsilon > 0$, compute positive numbers r, δ and a positive integer k such that $-y \in rC$, $\|x_k\| < \varepsilon / (n+r)$, and $\|x_k - z\| \geqslant \delta$ for all $z \in C$. Let

$$y_1 = y + (n+r)\, x_k.$$

Then

$$\|y - y_1\| = (n+r)\, \|x_k\| < \varepsilon.$$

On the other hand, since C is convex, for each $z \in nC$ we have

$$\frac{1}{n+r}\,(z-y) \in \frac{1}{n+r}\,(nC + rC) = C,$$

so

$$\|y_1 - z\| = (n+r) \left\| x_k - \frac{1}{n+r}\,(z-y) \right\| \geqslant (n+r)\,\delta.$$

Hence $y_1 \in -nC$. Since $\varepsilon > 0$ is arbitrary, it follows that $-nC$ is dense in X. Moreover, being a metric complement, $-nC$ is open. Applying the standard form of Baire's theorem, we now see that $\bigcap_{n\geqslant 1} -nC$ is inhabited, which is absurd since $X = \bigcup_{n\geqslant 1} nC$. We conclude that $0 \notin \overline{-C}$. □

Sometimes when we want to prove that a certain proposition P is absurd, we first prove that P implies LPO, and then (frequently by adapting a classical proof) show that the addition of LPO to our intuitionistic logic suffices for us to prove that P is false. The next lemma will enable us to rule out in this way an unwanted alternative in the proof of Banach's inverse mapping theorem.

Lemma 6.6.9. *Let f be a strongly extensional mapping of a complete metric space X into a metric space Y. Suppose that there exist $\alpha > 0$ and a sequence $(x_n)_{n\geqslant 1}$ converging to x in X such that $\rho(f(x), f(x_n)) > \alpha$ for each n. Then* LPO *holds.*

Proof. Given an increasing binary sequence $(\lambda_n)_{n\geqslant 1}$ with $\lambda_1 = 0$, construct a sequence $(z_n)_{n\geqslant 1}$ in X such that

- if $\lambda_k = 0$ for all $k \leqslant n$, then $z_n = x$, and
- if $\lambda_n = 1 - \lambda_{n-1}$, then $z_k = x_n$ for all $k \geqslant n$.

Then $(z_n)_{n\geqslant 1}$ is a Cauchy sequence in X and so converges to a limit $z \in X$. Either $f(x) \neq f(z)$ or else $\rho(f(x), f(z)) < \alpha$. In the first case, the strong extensionality of f shows that $x \neq z$; whence there exists n such that $x \neq z_n$ and therefore $\lambda_n = 1$. In the second case, if there exists n with $\lambda_n = 1 - \lambda_{n-1}$, then we obtain the contradiction $\rho(f(x), f(z)) = \rho(f(x), f(x_n)) > \alpha$; hence $\lambda_n = 0$ for all n. □

Lemma 6.6.10. *If* LPO *holds, then every separable subset of a metric space is located.*

Proof. Assuming LPO, consider a separable subset S of a metric space X. Let $(s_n)_{n\geqslant 1}$ be a dense sequence in S, let $x \in X$, and let $0 < \alpha < \beta$. Construct a binary sequence $(\lambda_n)_{n\geqslant 1}$ such that

$$\lambda_n = 0 \Longrightarrow \forall k \leqslant n \left(\rho(x, s_k) < \tfrac{1}{2}(\alpha + \beta)\right),$$
$$\lambda_n = 1 \Longrightarrow \exists k \leqslant n \left(\rho(x, s_k) > \alpha\right).$$

By LPO, either $\lambda_n = 0$ for all n, in which case, since $(s_n)_{n\geqslant 1}$ is dense in S, $\rho(x,s) \leqslant \frac{1}{2}(\alpha+\beta) < \beta$ for all $s \in S$; or else there exists n with $\lambda_n = 1$ and therefore $\rho(x, s_k) > \alpha$ for some $k \leqslant n$. It follows from the constructive least-upper-bound principle that $\rho(x, S)$ exists. □

This brings us to Ishihara's version of *Banach's inverse mapping theorem.*

Theorem 6.6.11. *Let T be a one-one, sequentially continuous linear mapping of a separable Banach space X onto a Banach space Y. Then T^{-1} is sequentially continuous.*

Proof. Let $(x_n)_{n\geqslant 1}$ be a sequence in X such that $Tx_n \longrightarrow 0$, and let $\varepsilon > 0$. By Corollary 3.1.7, the inverse linear mapping $T^{-1} : Y \longrightarrow X$ is strongly extensional; whence, by Lemma 3.2.2, either $\|x_n\| < \varepsilon$ for all sufficiently large n, or else $\|x_n\| > \varepsilon/2$ for infinitely many n. It suffices to rule out the latter case. To do so, we may assume that $\|x_n\| > \varepsilon/2$ for all n. By Lemma 6.6.9, LPO holds; so, by Lemma 6.6.10, every separable subset of Y is located. Let $(a_n)_{n\geqslant 1}$ be a dense sequence in $\overline{B}_X(0,1)$. Since T is sequentially continuous, $(Ta_n)_{n\geqslant 1}$ is dense in $T\left(\overline{B}_X(0,1)\right)$, which is therefore located. Writing $x_n' = 2\varepsilon^{-1}x_n$, we see that $\|x_n'\| > 1$ and that $Tx_n' \longrightarrow 0$ as $n \longrightarrow \infty$. Since T^{-1} is strongly extensional,

$$Tx_n' \in {\sim}T\left(\overline{B}_X(0,1)\right).$$

We now apply Lemma 6.6.2 with $y = T(2x_n')$, to produce $z_n \in X$ such that

$$\|Tz_n\| < \|Tx_n'\| + n^{-1} \quad \text{and} \quad \rho\left(T(2z_n), T\left(\overline{B}_X(0,1)\right)\right) > 0.$$

Then

$$\rho\left(Tz_n, T\left(\overline{B}_X(0,\tfrac{1}{2})\right)\right) > 0 \quad \text{and} \quad Tz_n \longrightarrow 0.$$

Hence

$$0 \in \overline{-T\left(B_X\left(0,\tfrac{1}{2}\right)\right)}.$$

This contradicts Lemma 6.6.8, since $T\left(B_X\left(0,\frac{1}{2}\right)\right)$ is both convex and absorbing. □

Recall that the *graph* of a mapping $T : X \longrightarrow Y$ is the set

$$\mathcal{G}(T) = \{(x, Tx) : x \in X\}.$$

As classically, Banach's inverse mapping theorem leads to a version of the *closed graph theorem*:

Corollary 6.6.12. *Let T be a linear mapping of a Banach space X into a Banach space Y such that $\mathcal{G}(T)$ is closed and separable. Then T is sequentially continuous.*

Proof. The mapping $p : (x, Tx) \rightsquigarrow x$ of the Banach space $\mathcal{G}(T)$ onto X is one-one and bounded linear. It follows from Theorem 6.6.11 that the inverse linear map is sequentially continuous, and hence that T is sequentially continuous. □

In the theory of unbounded operators, the graph plays a significant role. Particularly important properties for such a graph are closedness and locatedness. See [26, 83, 92] for more on such matters.

Exercises

1. Complete the proof of Proposition 6.1.1.

2. Complete the proof of Proposition 6.1.4.

3. Prove that a bounded linear mapping T of a normed space X into $\mathbb{C}^n$ is compact if and only if $f \circ T$ is normed for each linear functional f on $\mathbb{C}^n$. Prove that if also the sum of any two normed linear functionals on X is normed, then T is compact if and only if $p_k \circ T$ is normed for each k, where p_k denotes the mapping $(z_1, \ldots, z_n) \rightsquigarrow z_k$ on $\mathbb{C}^n$.

4. Let S be a compact operator on a Hilbert space H, and A a bounded operator on H. Prove that

   (a) λS is compact for each $\lambda \in \mathbb{C}$;

   (b) S^* exists and is compact;

   (c) AS is compact.

   Prove also that if A^* exists, then SA is compact.

5. A subset S of a normed space X is said the be *weakly totally bounded* if it is totally bounded relative to the locally convex structure defined on X by the seminorms $x \rightsquigarrow |f(x)|$, with f a normed linear functional on X. Prove the equivalence of the following conditions on X.

(a) The sum of any two normed linear functionals on X is normed.

(b) The unit ball of X is weakly totally bounded.

(*Hint:* To prove that (a) implies (b), note the second part of Exercise 3.)

6. An operator on a Hilbert space H is said to be *weakly compact* if it maps the unit ball of H to a weakly totally bounded subset of H. Prove that an operator T on H has an adjoint if and only if it is weakly compact.

7. Let $(e_n)_{n\geqslant 1}$ be an orthonormal basis of a separable Hilbert space H, and T an operator on H. Prove that the following conditions are equivalent:

   (a) T has an adjoint.

   (b) $\sum_{n=1}^{\infty} |\langle Te_n, e_k\rangle|^2$ converges for each k.

   (c) $\sum_{n=1}^{\infty} |\langle Te_n, y\rangle|^2$ converges for each $y \in H$.

8. Let T be a weak-sequentially open operator on a Hilbert space such that $\ker(T)$ is located. Prove that T is well-behaved (see Exercise 3 of Chapter 3).

9. Prove that every bounded linear mapping of a normed space onto a finite-dimensional Banach space is an open mapping.

10. Show that the statement "every normed linear mapping T of a Hilbert space onto a finite-dimensional Banach space is compact" implies an omniscience principle.

11. Show that the statement "every one-one compact linear mapping T of a Hilbert space into a finite-dimensional Banach space has finite-dimensional range" implies an omniscience principle.

12. Show that the existence of the norms of the operators can be removed from the hypotheses of Royden's version of the uniform boundedness theorem (Theorem 6.2.11).

13. Complete the details of the proof of Proposition 6.4.6.

14. A *partial isometry* is a jointed operator U on a Hilbert space H for which there exists a projection P, called the *initial projection* of U, such that $\|Ux\| = \|x\|$ for all $x \in \operatorname{ran}(P)$, and $Ux = 0$ for all $x \in \operatorname{ran}(P)^{\perp}$. Prove that $U^*U = P$, that UU^* is a projection, and that $\operatorname{ran}(UU^*) = \operatorname{ran}(U)$.

15. Let T be a bounded operator such that T^* exists and has located range, and let P be the projection of H on $\overline{\operatorname{ran}(T^*)}$. Prove that there exists a partial isometry

U on H whose initial projection is P such that $T = U|T|$ and $|T| = U^*T$. Prove also that UU^* is the projection of H on $\overline{\operatorname{ran}(T)}$. (The expression of T as $U|T|$ is called the *polar decomposition* of T, and is analogous to the modulus–argument form of a complex number.)

16. Prove that if a bounded jointed operator T on a Hilbert space H has a polar decomposition, then $\operatorname{ran}(T)$ is located.

17. Prove the converse of Proposition 6.5.12: if a linear mapping T between normed spaces maps Cauchy sequences to Cauchy sequences, then T is sequentially continuous.

18. Let E be a dense linear subspace of a normed space X, and T a sequentially continuous linear mapping of E into a Banach space Y. Show that T extends to a sequentially continuous linear mapping of X into Y.

19. Prove that if G is a compact generating set for a Banach space X, then there exists a compact generating set G' for X that is balanced and convex.

20. Let G be a balanced convex generating set for a normed space X. Prove that for each $x \in X$ and each $\varepsilon > 0$ there exist $t > 0$ and $g \in G$ such that $\|x - tg\| \leqslant \varepsilon$. Can we replace ε by 0 in this result?

Notes

Many of the results in this chapter come from papers by Ishihara and the authors. The work of Sections 1 and 2 is largely drawn from [61] and [63]. For related classical material see [86]. It is not known whether Theorems 6.2.1 and 6.2.2 can be extended to possibly unbounded convex subsets of a normed space. More general versions of Exercises 3–6 appeared in Ishihara's thesis [56].

Since we cannot guarantee that a bounded operator on a Hilbert space has an adjoint, when we discuss such matters as the Gelfand representation theorem for an operator algebra $\mathfrak{A}$, we need to postulate that $\mathfrak{A}$ is selfadjoint in the sense that each of its elements has an adjoint that also belongs to $\mathfrak{A}$; see [9, 12].

The principle BD-$\mathbb{N}$, introduced by Ishihara in [59], has the unusual feature of being provable classically, intuitionistically, recursively, but not, apparently, within BISH. It is an interesting problem—a part of *constructive reverse mathematics* [7, 62]—to identify classical theorems that are equivalent to BD-$\mathbb{N}$.

The usual classical proof of Lemma 6.4.7 is based on the nonconstructive monotone convergence theorem for sequences; see [47] (7.3.1.1). The constructive proof is much more informative, in that it provides the rate of convergence of the sequence of functions. A full constructive analysis of Dini's theorem is given in [8]; see also [21].

The existence of the square root of a selfadjoint operator is a special case of a more general result, the spectral theorem for sequences of commuting selfadjoint operators, which enables us to construct more general functions of an operator. Since that theorem requires measure theory, which we do not touch in this book, we refer the reader to the relevant chapters of [9] and [12].

It is easy to prove Lemma 6.5.3 classically, taking orthogonal complements in the identity $\ker(T^*T) = \ker(T^*)$.

For more on polar decompositions see [35].

Proposition 6.5.12 is trivial in CLASS, since in that context sequential continuity for linear maps implies uniform continuity.

With classical logic, Theorem 6.6.1 is a simple consequence of the standard form of Baire's theorem, the hypotheses of locatedness and the existence of $\rho(0, -C)$ being redundant. For other, constructively inequivalent, versions of Baire's theorem, see Chapter 2 of [34].

For more on open mapping theorems see [29] and [25]. It is interesting that sequential continuity, rather than boundedness, is the best we can get in the constructive versions of Banach's inverse mapping theorem and the closed graph theorem. An extension of the former is given in [64].

What can be said at all can be said clearly;
and whereof one cannot speak thereof one must be silent.

—Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*

References

1. O. Aberth, *Computable Analysis*, McGraw-Hill, New York, 1980.
2. O. Aberth, *Computable Calculus*, Academic Press, San Diego, CA, 2001.
3. P. Aczel and M. Rathjen, *Notes on Constructive Set Theory*, Report No. 40, Institut Mittag-Leffler, Royal Swedish Academy of Sciences, 2001.
4. G. Bachman and L. Narici, *Functional Analysis*, Academic Press, New York, 1966.
5. A. Bauer, Realizability as the connection between computable and constructive mathematics, Proceedings of CCA 2005, Kyoto, Japan, 25–29 August 2005; to appear.
6. M.J. Beeson, *Foundations of Constructive Mathematics*, Ergeb. Math. und ihrer Grenzgebiete, 3. Folge, **6**, Springer-Verlag, Heidelberg, 1985.
7. J. Berger and D.S. Bridges, A bizarre property equivalent to the Π_1^0-fan theorem, Logic J. of the IGPL, to appear.
8. J. Berger and P.M. Schuster, Classifying Dinis theorem, to appear in Notre Dame J. Formal Logic.
9. E.A. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967.
10. E.A. Bishop, Mathematics as a numerical language, in: *Intuitionism and Proof Theory* (J. Myhill, A. Kino, and R.E. Vesley, eds.), 53–71, North-Holland, Amsterdam, 1970.
11. E.A. Bishop, Schizophrenia in contemporary mathematics, in: *Errett Bishop: Reflections on Him and His Research* (M. Rosenblatt, ed.), 1–32, Contemporary Math. **39**, Amer. Math. Soc., Providence, RI, 1985.
12. E.A. Bishop and D.S. Bridges, *Constructive Analysis*, Grundlehren der math. Wissenschaften **279**, Springer-Verlag, Heidelberg, 1985.
13. E.A. Bishop and H. Cheng, *Constructive Measure Theory,* Mem. Amer. Math. Soc. **116**, 1972.
14. D.S. Bridges, On weak operator compactness of the unit ball of $\mathcal{L}(H)$, Zeit. math. Logik Grundlagen Math. **24**, 493–494, 1978.
15. D.S. Bridges, Recent progress in constructive approximation theory, in *The L.E.J. Brouwer Centenary Symposium* (A.S. Troelstra and D. van Dalen, eds.), 41–50, North–Holland, Amsterdam, 1982.
16. D.S. Bridges, A constructive Morse theory of sets, in *Mathematical Logic and Its Applications* (D. Skordev, ed.), 61–79, Plenum Publishing Corp., New York, 1987.
17. D.S. Bridges, Locatedness, convexity, and Lebesgue measurability, Quart. J. Math. Oxford (2) **39**, 411–421, 1988.

18. D.S. Bridges, A constructive look at the real number line, in: *Real Numbers, Generalizations of the Reals, and Theories of Continua* (P. Ehrlich, ed.), 29–92, Synthese Library **242**, Kluwer Academic Publishers, Dordrecht, 1994.
19. D.S. Bridges, Constructive mathematics: a foundation for computable analysis, Theoretical Computer Science **219**(1–2), 95–109, 1999.
20. D.S. Bridges, Constructive methods in mathematical economics, in *Mathematical Utility Theory*, J. Econ. (Zeitschrift für Nationalökonomie), Suppl. **8**, 1–21, 1999.
21. D.S. Bridges, Dinis Theorem: a constructive case study, in: *Combinatorics, Computability and Logic* (Proceedings of DMTCS01, Constanţa, Romania, 2–6 July 2001; C.S. Calude, M.J. Dinneen, S. Sburlan, eds.), 69–80, DMTCS Series **17**, Springer-Verlag, London, 2001.
22. D.S. Bridges, A weak constructive sequential compactness property and the fan theorem, Logic J. of the IGPL **13**(2), 151–158, 2005.
23. D.S. Bridges and N.F. Dudley Ward, Constructing ultraweakly continuous linear functionals on $\mathcal{B}(H)$, Proc. American Math. Soc. **126**(11), 3347–3353, 1998.
24. D.S. Bridges and H. Ishihara, Linear mappings are fairly well-behaved, Arch. Math. **54**, 558–562, 1990.
25. D.S. Bridges and H. Ishihara, A definitive constructive open mapping theorem?, Math. Logic Quarterly **44**, 545–552, 1998.
26. D.S. Bridges and H. Ishihara, Constructive unbounded operators, in: *Reuniting the Antipodes—Constructive and Nonstandard Views of the Continuum* (Proceedings of the Symposion in San Servolo/Venice, Italy, May 17–22, 1999; U. Berger, H. Osswald, and P.M. Schuster, eds.), 25–37, Synthese Library, Kluwer Academic Publishers, Dordrecht, December 2001.
27. D.S. Bridges, H. Ishihara, and P.M. Schuster, Sequential Compactness in Constructive Analysis, Österr. Akad. Wiss. Math.–Natur. Kl. Sitzungsber. II. **208**, 159–163, 1999.
28. D.S. Bridges, H. Ishihara, and L.S. Vîţă, Computing infima on convex sets, with applications in Hilbert spaces, Proc. Amer. Math. Soc. **132**(9)2723–2732, 2004.
29. D.S. Bridges, W.H. Julian, and R. Mines, A constructive treatment of open and unopen mapping theorems, Zeit. math. Logik Grundlagen Math. **35**, 29–43, 1989.
30. D.S. Bridges and A. Mahalanobis, Bounded variation implies regulated: a constructive proof, J. Symb. Logic **66**(4), 1695–1700, 2001.
31. D.S. Bridges and R. Mines, Sequentially continuous linear mappings in constructive analysis, J. Symb. Logic **63**(2), 579–583, 1998.
32. D.S. Bridges and G. Popa, Exact, continuous boundary crossings out of convex sets in $\mathbb{R}^N$, Quart. J. Math. **54**, 391–398, 2003.
33. D.S. Bridges and S. Reeves, Constructive mathematics, in theory and programming practice, Philosophia Mathematica **7**(1), 65–104, 1999.
34. D.S. Bridges and F. Richman, *Varieties of Constructive Mathematics*, London Math. Soc, Lecture Notes **97**, Cambridge Univ. Press, 1987.
35. D.S. Bridges, F. Richman, and P.M. Schuster, Adjoints, absolute values, and polar decompositions, J. Operator Theory **44**, 243–254, 2000.
36. D.S. Bridges, F. Richman, and Y. Wang, Sets, complements and boundaries, Proc. Koninklijke Nederlandse Akad. Wetenschappen (Indag. Math. N.S.) **7**(4), 425–445, 1996.
37. D.S. Bridges and L.S. Vîţă, Weak continuity properties in constructive analysis, Logic J. of the IGPL **7**(3), 277–281, 1999.
38. D.S. Bridges and L.S. Vîţă, Apartness spaces as a framework for constructive topology, Ann. Pure Appl. Logic **119** (1–3), 61–83, 2003.

39. D.S. Bridges and L.S. Vîţă, The constructive uniqueness of the locally convex topology on $\mathbb{R}^n$, in: *From Sets and Types to Topology and Analysis* (L. Crosilla and P.M. Schuster, eds.), 304–315, Oxford Logic Guides **32**, Clarendon Press, Oxford, 2005.
40. D.S. Bridges and L.S. Vîţă, *Apartness Spaces*, book, in preparation.
41. L.E.J. Brouwer, *Over de Grondslagen der Wiskunde,* doctoral thesis, University of Amsterdam, 1907. Reprinted with additional material (D. van Dalen, ed.) by Matematisch Centrum, Amsterdam, 1981.
42. R.L. Constable et al., *Implementing Mathematics with the Nuprl Proof Development System*, Prentice-Hall, Englewood Cliffs, New Jersey, 1986.
43. L. Crosilla and P.M. Schuster (eds.), *From Sets and Types to Topology and Analysis*, Oxford Logic Guides **32**, Clarendon Press, Oxford, 2005.
44. D. van Dalen (ed.), *Brouwers Cambridge Lectures on Intuitionism*, Cambridge University Press, Cambridge, 1981.
45. D. van Dalen, *Mystic, Geometer, and Intuitionist*, Clarendon Press, Oxford, 1999 (Vol. 1) and 2005 (Vol. 2).
46. R. Diaconescu, Axiom of choice and complementation, Proc. Amer. Math. Soc. **51**, 176–178, 1975.
47. J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York, 1960.
48. M.A.E. Dummett, *Elements of Intuitionism* (2nd edition), Oxford Logic Guides **39**, Clarendon Press, Oxford, 2000.
49. H.M. Edwards, *Essays in Constructive Mathematics*, Springer Science + Business Media, Inc., New York, 2005.
50. N. D. Goodman and J. Myhill, Choice implies excluded middle, Zeit. math. Logik und Grundlagen Math. **24**, 461.
51. S. Hayashi and H. Nakano, *PX: A Computational Logic*, MIT Press, Cambridge MA, 1988.
52. A. Heyting, Die formalen Regeln der intuitionistischen Logik, Sitzungsber. preuss. Akad. Wiss. Berlin, 42–56, 1930.
53. D. Hilbert, Über das Unendliche, Mathematische Annalen **95** , 161–190, 1926; translated in *Philosophy of Mathematics* (P. Benacerraf and H. Putnam, eds.), 183–201, Cambridge University Press, Cambridge, 1964.
54. D. Hilbert, Die Grundlagen der Mathematik, Abhandlungen Math. Sem. Univ. Hamburg **V**, 65–85, 1927.
55. H. Ishihara,On the constructive Hahn–Banach theorem, Bull. London Math. Soc. **21**, 79–81, 1989.
56. H. Ishihara, *Boundedness, normability and compactness of constructive linear mappings*, Ph.D. dissertation, Tokyo Institute of Technology, Tokyo, 1990.
57. H. Ishihara, Continuity and nondiscontinuity in constructive analysis, J. Symb. Logic **56**(4), 1349–1354, 1991.
58. H. Ishihara, Constructive compact operators on a Hilbert space, Ann. Pure Appl. Logic **52**, 31–37, 1991.
59. H. Ishihara, Continuity properties in constructive analysis, J. Symb. Logic **57**, 557–565, 1992.
60. H. Ishihara, A constructive version of Banachs inverse mapping theorem, New Zealand J. Math **23**, 71–75, 1994.
61. H. Ishihara, Locating subsets of a Hilbert space, Proc. Amer. Math. Soc. **129**(5), 1385–1390, 2001.
62. H. Ishihara, Constructive reverse mathematics: compactness properties, in: *From Sets and Types to Topology and Analysis* (L. Crosilla and P.M. Schuster, eds.), 245–267, Oxford Logic Guides **32**, Clarendon Press, Oxford, 2005.

63. H. Ishihara and L.S. Vîţă, Locating subsets of a normed space, Proc. Amer. Math. Soc. **131**(10), 3231–3239, 2003.
64. H. Ishihara and L.S. Vîţă, A constructive Banach inverse mapping theorem in $\mathcal{F}$-spaces, New Zealand J. Math., to appear.
65. R.V. Kadison and J.R. Ringrose, *Fundamentals of the Theory of Operator Algebras* (Vol. 1), Academic Press, New York, 1988.
66. B.A. Kushner, *Lectures on Constructive Mathematical Analysis*, Amer. Math. Soc., Providence RI, 1985.
67. A.A. Markov, *Theory of Algorithms* (Russian), Trudy Mat. Istituta imeni V.A. Steklova **42** (Izdatelstvo Akademi Nauk SSSR, Moskva), 1954; English translation by J.J. Schoor-Kan and PST staff, Israel Program for Scientific Translations, Jerusalem, 1961.
68. P. Martin-Löf, *Notes on Constructive Mathematics,* Almqvist and Wiksell, Stockholm, 1970.
69. P. Martin-Löf, An Intuitionistic Theory of Types: Predicative Part, in: *Logic Colloquium 1973* (H.E. Rose and J.C. Shepherdson, eds.), 73–118, North-Holland, Amsterdam, 1975.
70. P. Martin-Löf, Constructive mathematics and computer programming, in *Proc. 6th. Int. Congress for Logic, Methodology and Philosophy of Science* (L. Jonathan Cohen, ed.), 153–179, North-Holland, Amsterdam, 1980.
71. G. Metakides, A. Nerode, and R. Shore, Recursive limits on the Hahn–Banach theorem, in: *Errett Bishop: Reflections on Him and His Research* (M. Rosenblatt, ed.), 85–91, Contemporary Math. **39**, Amer. Math. Soc., Providence, R.I., 1985.
72. R. Mines, F. Richman, and W. Ruitenburg, *A Course in Constructive Algebra*, Universitext, Springer-Verlag, Heidelberg, 1988.
73. J.R. Moschovakis, The effect of Markovs principle on the intuitionistic continuum, preprint, UCLA, April 2005.
74. J. Myhill, Constructive set theory, J. Symb. Logic **40**, 347–382, 1975.
75. F. Richman (ed.), *Constructive Mathematics* (Proceedings of the Conference at Las Cruces, New Mexico, August 1980), Lecture Notes in Mathematics **873**, Springer-Verlag, Heidelberg, 1981.
76. F. Richman, The fundamental theorem of algebra: a constructive development without choice, Pacific J. Math. **196**, 213–230, 2000.
77. F. Richman, Adjoints and the image of the unit ball, Proc. Amer. Math. Soc. **129**, 1189–1193, 2001.
78. H.L. Royden, Aspects of constructive analysis, in: *Errett Bishop: Reflections on Him and His Research* (M. Rosenblatt, ed.), 57–64, Contemporary Mathematics **39**, American Math. Soc., 1985.
79. W. Rudin, *Real and Complex Analysis,* McGraw-Hill, New York, 1970.
80. E. Schechter, *Handbook of Analysis and Its Foundations*, Academic Press, San Diego, 1997.
81. P.M. Schuster, What is continuity, constructively?, J.UCS **11**(12), 2076–2085, 2005.
82. P.M Schuster, L.S. Vîţă, and D.S. Bridges, Apartness as a relation between subsets, in: *Combinatorics, Computability and Logic* (Proceedings of DMTCS01, Constanţa, Romania, 2–6 July 2001; C.S. Calude, M.J. Dinneen, S. Sburlan, eds.), 203–214, DMTCS Series **17**, Springer-Verlag, London, 2001.
83. B. Spitters, Located operators, Math. Logic Quarterly **48**(Suppl. 1), 107–122, 2002.
84. W.P. van Stigt, *Brouwers Intuitionism*, North–Holland, Amsterdam, 1990.
85. G. Stolzenberg, *Review of* [9], Bull. Amer. Math. Soc. **76**, 301–323, 1970.

86. W. Takahashi, *Nonlinear Functional Analysis,* Yokohama Publishers, 2000.
87. A. Takayama, *Mathematical Economics,* The Dryden Press, Hinsdale, IL, 1973.
88. A.S. Troelstra and D. van Dalen, *Constructivism in Mathematics: An Introduction* (two volumes), North Holland, Amsterdam, 1988.
89. F. Waaldijk, On the foundations of constructive mathematics, Foundations of Science **10**(3), 249–324, 2005.
90. H. Weber, Leopold Kronecker, Jahresber. der Deutschen Math. Verein **2**, 5–31, 1893.
91. K. Weihrauch, *Computable Analysis*, Springer-Verlag, Heidelberg, 2000.
92. F. Ye, Towards a constructive theory of unbounded operators, J. Symb. Logic **65**, 357–370, 2000.
93. R. Zach, Hilberts Program, in: *The Stanford Encyclopedia of Philosophy* (E.N. Zalta, ed.), URL `http://plato.stanford.edu/archives/fall2003/entries/hilbert-program/`.

The above list contains only a fraction of the publications on constructive mathematics that have appeared in the last forty years, and does not include the sources of all results in our book. The reader should not fall into the trap of believing that an unascribed result was first produced by the authors.

We mention two websites that may interest the reader:

`http://www.math.canterbury.ac.nz/php/groups/cm/faq/`
`http://plato.stanford.edu/entries/mathematics-constructive/`

In addition, many of the authors of items in the bibliography have websites that are worth a visit.

The primary historical reference on constructive analysis is [9], the review of which [85] is interesting in its own right. Later references for Bishop-style constructivism are [12, 34], the latter of which gives comparisons between BISH, INT, and RUSS. Beeson [6] and Troelstra–van Dalen [88] contain a wealth of information about the logic, philosophy, and practice of constructive mathematics. For some applications of constructive mathematics, see [20, 33, 92]. The definitive reference for constructive algebra is [72], but [49] should be consulted for more recent work in the field.

The classic work on intuitionism is [48]. The life and works of Brouwer himself are discussed in [44, 45, 84]. Martin-Löfs early work on constructive mathematics is found in [68], and his theory of types appears in [69].

Among the most recent varieties of computable analysis is that of Weihrauch [91]; the translation of BISH into Weihrauchs framework is described in [5].

Index

Universitext *(continued from p. ii)*

Hurwitz/Kritikos: Lectures on Number Theory
Jennings: Modern Geometry with Applications
Jones/Morris/Pearson: Abstract Algebra and Famous Impossibilities
Kac/Cheung: Quantum Calculus
Kannan/Krueger: Advanced Analysis
Kelly/Matthews: The Non-Euclidean Hyperbolic Plane
Kostrikin: Introduction to Algebra
Kuo: Introduction to Stochastic Integration
Kurzweil/Stellmacher: The Theory of Finite Groups: An Introduction
Lang: Introduction to Differentiable Manifolds
Lorenz: Algebra: Volume I: Fields and Galois Theory
Luecking/Rubel: Complex Analysis: A Functional Analysis Approach
MacLane/Moerdijk: Sheaves in Geometry and Logic
Marcus: Number Fields
Martinez: An Introduction to Semiclassical and Microlocal Analysis
Matsuki: Introduction to the Mori Program
McCarthy: Introduction to Arithmetical Functions
McCrimmon: A Taste of Jordan Algebras
Meyer: Essential Mathematics for Applied Fields
Mines/Richman/Ruitenburg: A Course in Constructive Algebra
Moise: Introductory Problems Course in Analysis and Topology
Morris: Introduction to Game Theory
Poizat: A Course In Model Theory: An Introduction to Contemporary Mathematical Logic
Polster: A Geometrical Picture Book
Porter/Woods: Extensions and Absolutes of Hausdorff Spaces
Procesi: Lie Groups
Radjavi/Rosenthal: Simultaneous Triangularization
Ramsay/Richtmyer: Introduction to Hyperbolic Geometry
Rautenberg: A Concise Introduction to Mathematical Logic, 2nd ed.
Reisel: Elementary Theory of Metric Spaces
Ribenboim: Classical Theory of Algebraic Numbers
Rickart: Natural Function Algebras
Rotman: Galois Theory
Rubel/Colliander: Entire and Meromorphic Functions
Runde: A Taste of Topology
Sagan: Space-Filling Curves
Samelson: Notes on Lie Algebras
Schiff: Normal Families
Shapiro: Composition Operators and Classical Function Theory
Simonnet: Measures and Probability
Smith: Power Series From a Computational Point of View
Smith/Kahanpää/Kekäläinen/Traves: An Invitation to Algebraic Geometry
Smorynski: Self-Reference and Modal Logic
Stillwell: Geometry of Surfaces
Stroock: An Introduction to the Theory of Large Deviations
Sunder: An Invitation to von Neumann Algebras
Tondeur: Foliations on Riemannian Manifolds
Toth: Finite Möbius Groups, Minimal Immersions of Spheres, and Moduli
Van Brunt: The Calculus of Variations
Weintraub: Galois Theory

Wong: Weyl Transforms
Zhang: Matrix Theory: Basic Results and Techniques
Zong: Sphere Packings
Zong: Strange Phenomena in Convex and Discrete Geometry