Mathematical Analysis

A Concise Introduction

Bernd S. W. Schröder
Louisiana Tech University
Program of Mathematics and Statistics
Ruston, LA

BICENTENNIAL 1807 WILEY 2007 BICENTENNIAL

WILEY-INTERSCIENCE
A John Wiley & Sons, Inc., Publication

This Page Intentionally Left Blank

Mathematical Analysis

The Wiley Bicentennial–Knowledge for Generations

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

William J. Pesce
President and Chief Executive Officer

Peter Booth Wiley
Chairman of the Board

Mathematical Analysis

A Concise Introduction

Bernd S. W. Schröder
Louisiana Tech University
Program of Mathematics and Statistics
Ruston, LA

WILEY-INTERSCIENCE
A John Wiley & Sons, Inc., Publication

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Wiley Bicentennial Logo: Richard J. Pacifico

Library of Congress Cataloging-in-Publication Data:

Schröder, Bernd S. W. (Bernd Siegfried Walter), 1966–
Mathematical analysis : a concise introduction / Bernd S.W. Schröder.
p. cm.
ISBN 978-0-470-10796-6 (cloth)
1. Mathematical analysis. I. Title.
QA300.S376 2007
515—dc22 2007024690

Printed in the United States of America.

Contents

Part II: Analysis in Abstract Spaces

Appendices

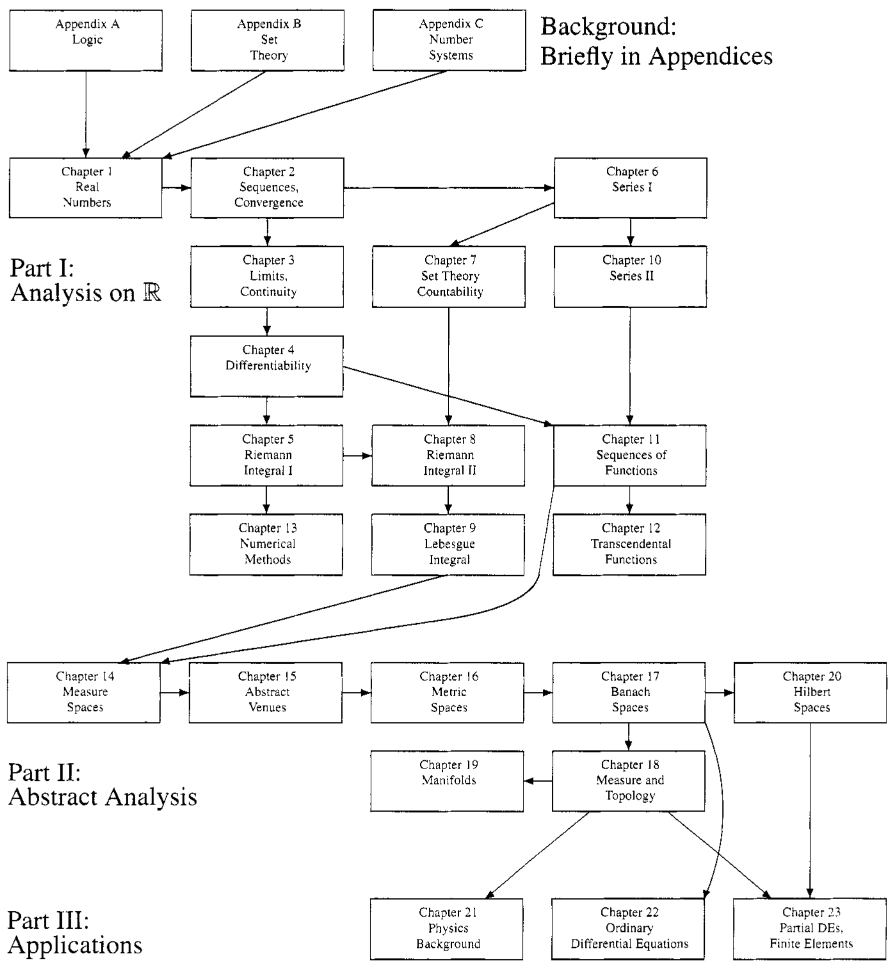

Figure 1: Content dependency chart with minimum prerequisites indicated by arrows. Some remarks, examples, and exercises in the later chapter might still depend on other earlier chapters, but this problem typically can be resolved by quoting a single result. Details about where and how the reader can "branch out" are given in **boxes** in the text.

Preface

This text is a self-contained introduction to the fundamentals of analysis. The only prerequisite is some experience with mathematical language and proofs. That is, it helps to be familiar with the structure of mathematical statements and with proof methods, such as direct proofs, proofs by contradiction, or induction. With some support in the right places, mostly in the early chapters, this text can also be used without prerequisites in a first proof class.

Mastering proofs in analysis is one of the key steps toward becoming a mathematician. To develop sound proof writing techniques, standard proof techniques are discussed early in the text and for a while they are pointed out explicitly. Throughout, proofs are presented with as much detail and as little hand waving as possible. This makes some proofs (for example, the density of $C[a, b]$ in $L^p[a, b]$ in Part II) notationally a bit complicated. With computers now being a regular tool in mathematics, the author considers this appropriate. When code is written for a problem, all details must be implemented, even those that are omitted in proofs. Seeing a few highly detailed proofs is reasonable preparation for such tasks. Moreover, to facilitate the transition to more abstract settings, such as measure, inner product, normed, and metric spaces, the results for single variable functions are proved using methods that translate to these abstract settings. For example, early proofs rely extensively on sequences and we also use the completeness of the real numbers rather than their order properties.

Analysis is important for applications, because it provides the abstract background that allows us to apply the full power of mathematics to scientific problems. This text shows that all abstractions are well motivated by the desire to build a strong theory that connects to specific applications. Readers who complete this text will be ready for *all* analysis-based and analysis-related subjects in mathematics, including complex analysis, differential equations, differential geometry, functional analysis, harmonic analysis, mathematical physics, measure theory, numerical analysis, partial differential equations, probability theory, and topology. Readers interested in motivation from physics are advised to browse Chapter 21, even if they have not read any of the earlier chapters.

Aside from the topics covered, readers interested in applications should note that the axiomatic approach of mathematics is similar to problem solving in other fields. In mathematics, theories are built on axioms. Similarly, in applications, models are subject to constraints. Neither the axioms, nor the constraints can be violated by the theory or model. Building a theory based on axioms fosters the reader's discipline to not make unwarranted assumptions.

Organization of the content. The text consists of three large parts. Part I, comprised of Chapters 1–13, presents the analysis of functions of one real variable, including a motivated introduction to the Lebesgue integral. Chapters 1–6 and 10–13 could be called "single variable calculus with proofs." For a smooth transition from calculus and a gradual increase in abstraction, Chapters 1–6 require very little set theory. Chapter 1 presents the properties of the real line and limits of sequences are introduced in Chapter 2. Chapters 3–5 present the fundamentals on continuity, differentiation, and (Riemann) integration in this order and Chapter 6 gives a first introduction to series.

Chapters 6–8 are motivated by the desire to further explore the Riemann integral while avoiding the excessive use of Riemann sums. This exploration is done with the Lebesgue criterion for Riemann integrability. Although this criterion requires the Lebesgue measure, the payoff is that many proofs become simpler. To quickly reach this criterion, the first presentation of series in Chapter 6 is deliberately kept short. It presents enough about series to allow the definition of Lebesgue measure. Chapter 7 presents fundamental notions of set theory. Most of these ideas are needed for Lebesgue measure, but, overall, Chapter 7 contains all the set theory needed in the remainder of the text. Chapter 8 finishes the presentation of the Riemann integral. With Lebesgue measure available, it is natural to investigate the Lebesgue integral in Chapter 9. This chapter could also be delayed to the end of Part I, but the author believes that early exposure to the crucial ideas will ease the later transition to measure spaces.

The analysis of single variable functions is finished with the rigorous introduction of the transcendental functions. The necessary background on power series is explored in Chapter 10. Chapter 11 presents some fundamentals on the convergence of sequences of functions and Chapter 12 is devoted to the transcendental functions themselves. Chapter 13 discusses general numerical methods, but transcendental functions provide a rich test bed for the methods presented.

Part I of the text can be read or presented in many orders. Figure 1 shows the prerequisite structure of the text. Prerequisites for each chapter have deliberately been kept minimal. In this fashion, the order of topics in the reader's first contact with proofs in analysis can be adapted to many readers' preferences. Most notably, the intentionally early presentation of Lebesgue integration can be postponed to the end of Part I if so desired. Throughout, the author intends to keep the reader engaged by providing motivation for all abstractions. Consequently, as Figure 1 and the table of contents indicate, some concepts and results are presented in a "just-in-time" fashion rather than in what may be considered their traditional place. If a concept is needed in an exercise before the concept is "officially" defined in the text, the concept will be defined in the exercise and in the text.

Part II, comprised of Chapters 14–20, explores how the appropriate abstractions lead to a powerful and widely applicable theoretical foundation for all branches of applied mathematics. The desire to define an integral in d-dimensional space provides a natural motivation to introduce measure spaces in Chapter 14. This chapter facilitates the transition to more abstract mathematics by frequently referring back to corresponding results for the one dimensional Lebesgue integral. The proofs of these results usually are verbatim the same as in the one-dimensional setting. Moreover, this early introduction makes L^p spaces available as examples for the rest of the text. The abstract venues of analysis are then presented in Chapter 15, which provides all examples for the rest of Part II.

The fundamentals on metric spaces and continuity are presented in Chapter 16. As with measure spaces, for several results on metric spaces the reader is referred back to the corresponding proof for single variable functions. Proofs are no longer verbatim the same and abstraction is facilitated by translating proofs from a familiar setting to the new setting while analyzing similarities and differences. In a class, the author suggests that the teacher fill in some of these proofs to demonstrate the process.

Chapter 17 presents the fundamentals on normed spaces and differentiation. Again, ideas are similar to those for functions of a single variable, but this time the abstraction goes beyond translation. With all three fundamental concepts (integration, continuity, and differentiation) available in the abstract setting, Chapter 18 shows the interrelationship between concepts presented separately before, culminating in the Multivariable Substitution Formula.

The second part is completed by a presentation of the fundamentals of analysis on manifolds, together with a physical interpretation of key concepts in Chapter 19 and by an introduction to Hilbert spaces in Chapter 20.

The remaining chapters give a brief outlook to applied subjects in which analysis is used, specifically, physics in Chapter 21, ordinary differential equations in Chapter 22, and partial differential equations and the finite element method in Chapter 23. Each of these chapters can only give a taste of its subject and I encourage the reader to go deeper into the utterly fascinating applications that lie behind part III. The mathematical preparation through this text should facilitate the transition.

It should be possible to cover the bulk of the text in a two course sequence. Although Chapters 14-16 should be read in order, depending on the available time, the pace and the choice of topics, any of Chapters 17-23 can serve as a capstone experience.

How to read this text. Mathematics in general, and analysis in particular, is not a spectator sport. It is learned by doing. To allow the reader to "do" mathematics, each section has exercises of varying degrees of difficulty. Some exercises require the adaptation of an argument in the text. These exercises are also intended to make the reader critically analyze the argument before adapting it. This is the first step toward being able to write proofs. Of course the need for very critical (and slow) reading of mathematics is nicely summed up in the old quote that "To read without a pencil is daydreaming." The reader should ask him/herself after every sentence "What does this mean? Why is this justified?" Making notes in the margin to explain the harder steps will allow the reader to answer these questions more easily in the second and third readings of a proof. So it is important to read thoroughly and slowly, to make notes and to reread as often as needed. The extensive index should help with unknown or forgotten terminology as necessary. Other exercises have hints on how to create a proof that the reader has not seen before. These exercises require the use of proof techniques in a new setting. Finally, there are also exercises without hints. Being able to create the proof with nothing but the result given is the deepest task in a mathematics course. This is not to say that exercises without hints are always the hardest and adaptations are always the easiest, but in many cases this is true. Finally, some exercises give a sequence of hints and intermediate results leading up to a famous theorem or a specific example. These exercises could also be used as mini-projects. In a class, some of them could be the basis for separate lectures that spotlight a particular theorem or example.

To get the most out of this text, the reader is encouraged to *not* look for hints and solutions in other background materials. In fact, even for proofs that are adaptations of proofs in this text, it is advantageous to try to create the proof *without* looking up the proof that is to be adapted. There is evidence that the struggle to solve a problem, which can take days for a single proof, is exactly what ultimately contributes to the development of strong skills. "Shortcuts," while pleasant, can actually diminish this development. Readers interested in quantitative evidence that shows how the struggle to acquire a skill actually can lead to deeper learning may find the article [4] quite enlightening. A better survival mechanism than shortcuts is the development of connections between newly learned content and existing knowledge. The reader will need to find these connections to his/her existing knowledge, but the structure of the text is intended to help by motivating all abstractions. Readers interested in how knowledge is activated more easily when it was learned in a known context may be interested in the article [5].

Acknowledgments. Strange as it may sound, I started writing this text in the spring of 1987, as I prepared for my oral final examination in the traditional Analysis I–III sequence in Germany. Basically, I took all topics in the sequence and arranged them in what was the most logical fashion to me at the time. Of course, these notes are, in retrospect, immature. But they did a lot to shape my abilities and they were a good source of ideas and exercises. In this respect, I am indebted to my teachers for this sequence: Professor Wegener and teaching assistant Ms. Lange for Analysis I, Professor Kutzler and teaching assistant *Herr* Böttger for Analysis II-III as well as Professor Herz in whose Differential Equations class I first saw analysis "at work." With all due respect to the other individuals, to me and many of my fellow students, the force that drove us in analysis (and beyond) was Herr Böttger. This gentleman was uncompromising in his pursuit of mathematical excellence and we feared as well as looked forward to his demanding exercise sets. He was highly respected because he was ready to spend hours with anyone who wanted to talk mathematics. Those who kept up with him were extremely well prepared for their mathematical careers. Incidentally, Dr. Ansgar Jüngel, whose notes I used for the chapter on the finite element method, took the above mentioned classes with me. The thorough preparation through these classes is the main reason why most of this text was comparatively easy to write. If this text does half as good a job as Herr Böttger did with us, it has more than achieved its purpose.

It was thrilling to test my limitations, it was humbling to find them and ultimately I was left awed once more by the beauty of mathematics. When my abilities were insufficient to proceed, I used the texts listed in the bibliography for proofs, hints or to structure the presentation. To make the reader fully concentrate on matters at hand, and to force myself to make the exposition self-contained, outside references are limited to places where results were beyond the scope of this exposition. A solid foundation will allow readers to judiciously pick their own resources for further study. Nonetheless, it is appropriate to recognize the influence of the works of a number of outstanding individuals. I used Adams [2], Renardy and Rogers [23], Yosida [33] and Zeidler [34] for Sobolev spaces, Aris [3], Cramer's `http://www.navier-stokes.net/`, and

Welty, Wicks and Wilson [31] for fluid dynamics, Chapman [6] for heat transfer, Cohn [7] for measure theory, Dieudonné [8] for differentiation in Banach spaces, Dodge [9] and Halmos [13] for set theory, Ferguson [10], Sandefur [24] and Stoer and Bulirsch [28] for numerical analysis, Halliday, Resnick and Walker [12] for elementary physics, Hewitt and Stromberg [14], Heuser [15], [16], Johnsonbaugh and Pfaffenberger [20], Lehn [22] and Stromberg [29] for general background on analysis, Heuser [17] for functional analysis, Hurd and Loeb [18] for the use of quantifiers in logic, Jüngel [21] and Šolín [25] for the finite element method, Spivak [26], [27] for manifolds, Torchinsky [30] for Fourier series, Willard [32] for topology, and the Online Encyclopaedia of Mathematics `http://eom.springer.de/` for quick checks of notation and definitions. Readers interested in further study of these subjects may wish to start with the above references.

The first draft of the manuscript was used in my analysis classes in the Winter and Spring quarters of 2007. The first class covered Chapters 1–9, the second covered Chapters 11 and 14–18 (with some strategic "fast forwards"). This setup assured that graduating students would have full exposure to the essentials of analysis on the real line and to as much abstract analysis as possible without "handwaving arguments." I am grateful to the students in these classes for keeping up with the pace, solving large numbers of homework problems, being patient with the typos we found and also for suggesting at least one order in which to present the material that I had not considered. The students' evaluations (my best ever) also reaffirmed for me that people will enjoy, or at least accept and honor, a challenge, and that an ambitious, motivated course should be the way to go. Devery Rowland once more did an excellent job printing drafts of the text for the classes.

Aside from the referees, several colleagues also commented on this text and I owe them my thanks for making it a better product. In particular, I would like to thank Natalia Zotov for some comments on an early version that significantly improved the presentation, and Ansgar Jüngel for pointing out some key references on Sobolev spaces. Although I hope that we have found all remaining errors and typos, any that remain are my responsibility and mine alone. I request readers to report errors and typos to me so I can post an errata. My contacts at Wiley, Susanne Steitz, Jacqueline Palmieri, and Melissa Yanuzzi bore with me when the stress level rose and their patience made the publishing process very smooth.

As always, this work would not have been possible without the love of my family. It is truly wonderful to be supported by individuals who accept your decision to spend large amounts of time reliving your formative years.

Finally, I was sad to learn that Herr Böttger died unexpectedly a few years after I had my last class with him. Sir, this one's for you.

Ruston, LA, August 30, 2007

Bernd Schröder

Part I

Analysis of Functions of a Single Real Variable

Chapter 1

The Real Numbers

This investigation of analysis starts with minimal prerequisites. Regarding set theory, the terms "set" and "element" will remain undefined, as is customary in mathematics to avoid paradoxes. The **empty set** $\emptyset$ is the set that has no elements. The statement "$e \in S$" says that e is an element of the set S. The statement "$A \subseteq B$" says that every element of A is an element of B. Sets A and B are equal if and only if $A \subseteq B$ and $B \subseteq A$. The statement "$A \subset B$" says that $A \subseteq B$ and $A \neq B$. Subsets will be defined as "$A = \{x \in S : \langle\text{property}\rangle\}$," that is, with a statement from which set S the elements of A are taken and a property describing them. The **union** of two sets A and B is $A \cup B = \{x : x \in A \text{ or } x \in B\}$, the **intersection** is $A \cap B = \{x : x \in A \text{ and } x \in B\}$. Union and intersection of finitely many sets are denoted $\bigcup_{j=1}^{n} A_j$ and $\bigcap_{j=1}^{n} A_j$, respectively, and the **relative complement** of B in A is $A \setminus B = \{x \in A : x \notin B\}$. Further details on set theory are purposely delayed until Section 7.1. Until then, we focus on analytical techniques. Any required notions of set theory will be clarified on the spot.

To define properties, sometimes the **universal quantifier** "$\forall$" (read "for all") or the **existential quantifier** "$\exists$" (read "there exists") are used. Formal logic is described in more detail in Appendix A. Finally, the reader needs an intuitive idea what a function, a relation and a binary operation are. Details are relegated to Appendices B.2 and C.2.

The real numbers $\mathbb{R}$ are the "staging ground" for analysis. They can be characterized as the unique (up to isomorphism) mathematical entity that satisfies Axioms 1.1, 1.6, and 1.19. That is, they are the unique linearly ordered, complete field (see Exercise 1-30). In this chapter, we introduce the axioms for the real numbers and some fundamental consequences. These results assure that the real numbers indeed have the properties that we are familiar with from algebra and calculus.

1.1 Field Axioms

The description of the real numbers starts with their algebraic properties.

Axiom 1.1 *The* **real numbers** $\mathbb{R}$ *are a* **field**. *That is,* $\mathbb{R}$ *has at least two elements and there are two binary operations,* **addition** $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *and* **multiplication** $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, *so that*

1. *Addition is* **associative**, *that is, for all* $x, y, z \in \mathbb{R}$ *we have*
   $(x + y) + z = x + (y + z)$.
2. *Addition is* **commutative**, *that is, for all* $x, y \in \mathbb{R}$ *we have*
   $x + y = y + x$.
3. *There is a* **neutral element** 0 *for addition, that is, there is an element* $0 \in \mathbb{R}$ *so that for all* $x \in \mathbb{R}$ *we have* $x + 0 = x$.
4. *For every element* $x \in \mathbb{R}$ *there is an* **additive inverse** *element* $(-x)$ *so that*
   $x + (-x) = 0$.
5. *Multiplication is* **associative**, *that is, for all* $x, y, z \in \mathbb{R}$ *we have*
   $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.
6. *Multiplication is* **commutative**, *that is, for all* $x, y \in \mathbb{R}$ *we have*
   $x \cdot y = y \cdot x$.
7. *There is a* **neutral element** 1 *for multiplication, that is, there is an element* $1 \in \mathbb{R}$ *so that for all* $x \in \mathbb{R}$ *we have* $1 \cdot x = x$.
8. *For every element* $x \in \mathbb{R} \setminus \{0\}$ *there is a* **multiplicative inverse** *element* x^{-1} *so that* $x \cdot x^{-1} = 1$.
9. *Multiplication is* **(left) distributive** *over addition, that is, for all* $\alpha, x, y \in \mathbb{R}$ *we have* $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$.

As is customary for multiplication, the dot between factors is usually omitted.

Fields are investigated in detail in abstract algebra. For analysis, it is most effective to remember that the field axioms guarantee the properties needed so that we can perform algebra and arithmetic "as usual." Some of these properties are exhibited in this section and in the exercises. The exercises also include examples that show that not every field needs to be infinite (see Exercises 1-7–1-9).

Theorem 1.2 *The following are true in* $\mathbb{R}$:

1. *For all* $x \in \mathbb{R}$, *we have* $0x = 0$.
2. $0 \neq 1$.
3. *Additive inverses are unique. That is, if* $x \in \mathbb{R}$ *and* x' *and* $\overline{x}$ *both have the property in part 4 of Axiom 1.1, then* $x' = \overline{x}$.
4. *For all* $x \in \mathbb{R}$, *we have* $(-1)x = -x$.

Proof. *Early in the text, proofs will sometimes be interrupted by comments in italics to point out standard formulations and proof techniques.*

To prove part 1, let $x \in \mathbb{R}$. Then the axioms allow us to obtain the following equation. $0x \overset{\text{Ax.3}}{=} (0+0)x \overset{\text{Ax.6}}{=} x(0+0) \overset{\text{Ax.9}}{=} x0 + x0 \overset{\text{Ax.6}}{=} 0x + 0x$. This implies

$$0 \overset{\text{Ax.4}}{=} 0x + (-0x) \overset{\text{above}}{=} (0x + 0x) + (-0x) \overset{\text{Ax.1}}{=} 0x + \big(0x + (-0x)\big) \overset{\text{Ax.4}}{=} 0x + 0 \overset{\text{Ax.3}}{=} 0x$$

as was claimed. *The proof of part 1 shows how every step in a proof needs to be justified. Usually we will not explicitly justify each step in a computation with an axiom or a previous result. However, the reader should always mentally fill in the justification. The practice of filling in these justifications should be started in the computations in the remainder of this proof.*

To prove part 2, first note that, because $\mathbb{R}$ has at least two elements, there is an $x \in \mathbb{R} \setminus \{0\}$. Now suppose for a contradiction *(see Standard Proof Technique 1.4 below)* that $0 = 1$. Then $x = 1 \cdot x = 0 \cdot x = 0$ is a contradiction to $x \in \mathbb{R} \setminus \{0\}$.

For part 3, note that if x' and $\overline{x}$ both have the property in part 4 of Axiom 1.1, then $x' = x' + 0 = x' + (x + \overline{x}) = \big(x' + x\big) + \overline{x} = \big(x + x'\big) + \overline{x} = 0 + \overline{x} = \overline{x} + 0 = \overline{x}$. *Note that the statement of part 3 already encodes the typical approach to a uniqueness proof (see Standard Proof Technique 1.5 below).*

Finally, for part 4 note that $x + (-1)x = 1x + (-1)x = \big(1 + (-1)\big)x = 0x = 0$. Because by part 3 additive inverses are unique, $(-1)x$ must be the additive inverse $-x$ of x. *The last step is a typical application of modus ponens, see Standard Proof Technique 1.3 below.* ■

To familiarize the reader with standard proof techniques, these techniques will be pointed out explicitly in the early part of the text. The techniques presented in Chapter 1 are general proof techniques applicable throughout mathematics. Techniques presented in later chapters are mostly specific to analysis.

Standard Proof Technique 1.3 The simplest mathematical proof technique is a **direct proof** in which a result that says "A implies B" is applied after we have proved that A is true. Truth of A and of "A implies B" guarantees truth of B. This technique is also called **modus ponens**. An example is in the proof of part 4 of Theorem 1.2. □

Standard Proof Technique 1.4 In a proof by **contradiction**, we suppose the contrary (the negation, also see Appendix A.2) of what is claimed is true and then we derive a contradiction. Typically, we derive a statement and its negation, which is a contradiction, because they cannot both be true. For an example, see the proof of part 2 of Theorem 1.2 above. Given that the reasoning that led to the contradiction is correct, the contradiction must be caused by the assumption that the contrary of the claim is true. Hence, the contrary of the claim must be false, because true statements cannot imply false statements like contradictions (see part 3 of Definition A.2 in Appendix A). But this means the claim must be true.

We will usually indicate proofs by contradiction with a starting statement like "suppose for a contradiction." □

Standard Proof Technique 1.5 For many mathematical objects it is important to assure that they are the *only* object that has certain properties. That is, we want to assure that the object is unique. In a typical **uniqueness proof**, we assume that there is more than one object with the properties under investigation and we prove that any two of these objects must be equal. Part 3 of Theorem 1.2 shows this approach. □

Exercises

1-1. Prove that $(-1) \cdot (-1) = 1$.

1-2. $\cdot$ **is right distributive over** $+$. Prove that for all $x, y, z \in \mathbb{R}$ we have $(x + y)z = xz + yz$.

1-3. **Multiplicative inverses are unique**. Prove that if $x \in \mathbb{R}$ and x' and $\overline{x}$ both have the property in part 8 of Axiom 1.1 then $x' = \overline{x}$.

1-4. Prove that 0 does not have a multiplicative inverse.

1-5. Prove that if $x, y \neq 0$, then $(xy)^{-1} = y^{-1}x^{-1}$. Conclude in particular that $xy \neq 0$.

1-6. Prove each of the **binomial formulas** below. Justify each step with the appropriate axiom.

(a) $(a + b)^2 = a^2 + 2ab + b^2$ (b) $(a - b)^2 = a^2 - 2ab + b^2$

(c) $(a + b)(a - b) = a^2 - b^2$

1-7. Prove that the set $\{0, 1\}$ with the usual multiplication and the usual addition, except that $1 + 1 := 0$, is a field. That is, prove that the set and addition and multiplication as stated have the properties listed in Axiom 1.1.

1-8. Prove that the set $\{0, 1, 2\}$ with the sum and product of two elements being the remainder obtained when dividing the regular sum and product by 3 is a field.

1-9. A property and some finite fields.

(a) Let F be a field and let $x, y \in F$. Prove that $x \cdot y = 0$ if and only if $x = 0$ or $y = 0$.

(b) Prove that the set $\{0, 1, 2, 3\}$ with the sum and product of two elements being the remainder obtained when dividing the regular sum and product by 4 is *not* a field.

(c) Prove that the set $\{0, 1, \ldots, p - 1\}$ with the sum and product of two elements being the remainder obtained when dividing the regular sum and product by p is a field if and only if p is a prime number.

1.2 Order Axioms

Exercises 1-7–1-9c show that the field axioms alone are not enough to describe the real numbers. In fact, fields need not even be infinite. However, aside from executing the familiar algebraic operations, we can also compare real numbers. This section presents the order relation on the real numbers and its properties.

Axiom 1.6 *The real numbers* $\mathbb{R}$ *contain a subset* $\mathbb{R}^+$, *called the* **positive real numbers** *such that*

1. *For all* $x, y \in \mathbb{R}^+$, *we have* $x + y \in \mathbb{R}^+$ *and* $xy \in \mathbb{R}^+$,

2. *For all* $x \in \mathbb{R}$, *exactly one of the following three properties holds.*
   Either $x \in \mathbb{R}^+$ *or* $-x \in \mathbb{R}^+$ *or* $x = 0$.

A real number x is called **negative** if and only if $-x \in \mathbb{R}^+$.

Once positive numbers are defined, we can define an order relation. As usual, instead of writing $y + (-x)$ we write $y - x$ and call it the **difference** of x and y. The binary operation "$-$" is called **subtraction**.

The phrase "if and only if," which is used in definitions and biconditionals, is normally abbreviated with the artificial word "**iff**."

Definition 1.7 *For $x, y \in \mathbb{R}$, we say x is* **less than** *y, in symbols $x < y$, iff $y - x \in \mathbb{R}^+$. We say x is* **less than or equal to** *y, denoted $x \leq y$, iff $x < y$ or $x = y$. Finally, we say x is* **greater than** *y, denoted $x > y$, iff $y < x$, and we say x is* **greater than or equal to** *y, denoted $x \geq y$, iff $y \leq x$.*

The relation $\leq$ satisfies the properties that define an order relation.

Proposition 1.8 *The relation $\leq$ is an* **order** *relation on $\mathbb{R}$. That is,*

1. *$\leq$ is* **reflexive**. *For all $x \in \mathbb{R}$ we have $x \leq x$,*
2. *$\leq$ is* **antisymmetric**. *For all $x, y \in \mathbb{R}$ we have that $x \leq y$ and $y \leq x$ implies $x = y$,*
3. *$\leq$ is* **transitive**. *For all $x, y, z \in \mathbb{R}$, we have that $x \leq y$ and $y \leq z$ implies $x \leq z$.*

Moreover, the relation $\leq$ is a **total order** *relation, that is, for any two $x, y \in \mathbb{R}$ we have that $x \leq y$ or $y \leq x$.*

Proof. The relation $\leq$ is reflexive, because it includes equality.

For antisymmetry, let $x \leq y$ and $y \leq x$ and suppose for a contradiction that $x \neq y$. Then $x - y \in \mathbb{R}^+$ and $-(x - y) = y - x \in \mathbb{R}^+$, which cannot be by Axiom 1.6. Thus $\leq$ must be antisymmetric.

For transitivity, let $x \leq y$ and $y \leq z$. There is nothing to prove if one of the inequalities is an equality. Thus we can assume that $x < y$ and $y < z$, which means $y - x \in \mathbb{R}^+$ and $z - y \in \mathbb{R}^+$. But then $\mathbb{R}^+$ contains $(z - y) + (y - x) = z - x$, and hence $x < z$. We have shown that for all $x, y, z \in \mathbb{R}$ the inequalities $x \leq y$ and $y \leq z$ imply $x \leq z$, which means that $\leq$ is transitive.

For the "moreover" part note that if $x, y \in \mathbb{R}$, then $y - x \in \mathbb{R}$ and we have either $y - x \in \mathbb{R}^+$, which means $x < y$, or $y - x = 0$, which means $y = x$, or $x - y = -(y - x) \in \mathbb{R}^+$, which means $y < x$. Therefore for all $x, y \in \mathbb{R}$ one of $x \leq y$ or $y \leq x$ holds, and hence $\leq$ is a total order. ■

Once an order relation is established, we can define intervals.

Definition 1.9 *An* **interval** *is a set $I \subseteq \mathbb{R}$ so that for all $c, d \in I$ and $x \in \mathbb{R}$ the inequalities $c < x < d$ imply $x \in I$. In particular for $a, b \in \mathbb{R}$ with $a < b$ we define*

1. $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$,
2. $(a, b) := \{x \in \mathbb{R} : a < x < b\}$, $(a, \infty) := \{x \in \mathbb{R} : a < x\}$, $(-\infty, b) := \{x \in \mathbb{R} : x < b\}$, $(-\infty, \infty) := \mathbb{R}$,

3. $[a, b) := \{x \in \mathbb{R} : a \le x < b\}$, $[a, \infty) := \{x \in \mathbb{R} : a \le x\}$,

4. $(a, b] := \{x \in \mathbb{R} : a < x \le b\}$, $(-\infty, b] := \{x \in \mathbb{R} : x \le b\}$.

The points a and b are also called the **endpoints** *of the interval. An interval that does not contain either of its endpoints (where $\pm\infty$ are also considered to be "endpoints") is called* **open**. *An interval that contains exactly one of its endpoints is called* **half-open** *and an interval that contains both its endpoints is called* **closed**.

For the first part of this text, the domains of functions will almost exclusively be intervals. Because analysis requires extensive work with inequalities, we need to investigate how the order relation relates to the algebraic operations.

Theorem 1.10 *Properties of the order relation. Let $x, y, z \in \mathbb{R}$.*

1. *The number x is positive iff $x > 0$ and x is negative iff $x < 0$.*

2. *If $x \le y$, then $x + z \le y + z$.*

3. *If $x \le y$ and $z > 0$, then $xz \le yz$.*

4. *If $x \le y$ and $z < 0$, then $xz \ge yz$.*

5. *If $0 < x \le y$, then $y^{-1} \le x^{-1}$.*

Similar results can be proved for other combinations of strict and nonstrict inequalities. We will not state these here, but instead trust that the reader can make the requisite translation from the statements in this theorem.

Proof. Parts 1 and 2 are left to the reader as Exercises 1-10a and 1-10b. *Throughout this text, parts of proofs will be delegated to the reader to facilitate a better connection to the material presented.*

For part 3, let $x \le y$ and let $z > 0$. Then, $y - x \in \mathbb{R}^+$ or $y = x$. In case $y = x$, we obtain $yz = xz$ and thus, in particular, $xz \le yz$. In case $y - x \in \mathbb{R}^+$, note that $z > 0$ means $z \in \mathbb{R}^+$, and hence $yz - xz = (y - x)z \in \mathbb{R}^+$. By definition, this implies $xz < yz$, and in particular $xz \le yz$. Because we have shown $xz \le yz$ in each case, the result is established. *All proofs in this section are done with the above kind of case distinction (see Standard Proof Technique 1.11).*

For part 4, let $x \le y$ and let $z < 0$. Then, $y - x \in \mathbb{R}^+$ or $y = x$. In case $y = x$, we obtain $yz = xz$, and hence $xz \ge yz$. In case $y - x \in \mathbb{R}^+$, note that $z < 0$ means $-z \in \mathbb{R}^+$, and hence $xz - yz = (x - y)z = (y - x)(-z) \in \mathbb{R}^+$. By definition, this implies $yz < xz$, and hence $yz \le xz$, which establishes the result.

For part 5, first note that there is nothing to prove if $x = y$. Hence, we can assume that $x < y$. Suppose for a contradiction that $x^{-1} < y^{-1}$. Then by part 3 we have that $1 = x^{-1}x < y^{-1}x$, and hence $x < y \cdot 1 < yy^{-1}x = x$, contradiction. ■

Standard Proof Technique 1.11 When several possibilities must be considered in a proof, the proof usually continues with separate arguments for each possibility. The proof is complete when each separate argument has led to the desired conclusion. This type of proof is also called a proof by **case distinction**. □

We conclude this section by introducing the absolute value function and some of its properties.

Definition 1.12 *For $x \in \mathbb{R}$, we set $|x| = \begin{cases} x; & \text{if } x \geq 0, \\ -x; & \text{if } x < 0, \end{cases}$ and we call it the* **absolute value** *of x.*

Theorem 1.13 summarizes the properties of the absolute value. The numbering is adjusted so that properties 1, 2, and 3 correspond to the analogous properties for norms (see Definition 15.38). *We will formulate many results in the first part of the text to be analogous or easily generalizable to more abstract settings, but we will usually do so without explicit forward references. In this fashion many abstract situations will be more familiar because of similarities to situations investigated in the first part.*

Theorem 1.13 *Properties of the absolute value.*

0. *For all $x \in \mathbb{R}$, we have $|x| \geq 0$,*
1. *For all $x \in \mathbb{R}$, we have $|x| = 0$ iff $x = 0$,*
2. *For all $x, y \in \mathbb{R}$, we have $|xy| = |x||y|$,*
3. **Triangular inequality**. *For all $x, y \in \mathbb{R}$, we have $|x + y| \leq |x| + |y|$.*
4. **Reverse triangular inequality**. *For all $x, y \in \mathbb{R}$, we have $\big||x| - |y|\big| \leq |x - y|$.*

Proof. For part 0, let $x \in \mathbb{R}$. In case $x \geq 0$, by Definition 1.12 we have $|x| = x \geq 0$. In case $x < 0$, we have $x \notin \mathbb{R}^+$ and by part 2 of Axiom 1.6 we conclude $-x > 0$. Because in this case $|x| = -x > 0$, part 0 follows.

Throughout the text, the two implications of a biconditional "A iff B" will be referred to as "$\Rightarrow$," denoting "if A, then B" and "$\Leftarrow$," denoting "if B, then A."

For part 1, note that the direction "$\Leftarrow$" is trivial, because $|0| = 0$. For the direction "$\Rightarrow$," let $x \in \mathbb{R}$ be so that $|x| = 0$ and suppose for a contradiction that $x \neq 0$. If $x > 0$, then $0 < x = |x| = 0$, a contradiction. *(Note that the previous sentence is a short proof by contradiction that is part of a longer proof by contradiction.)* Therefore $x < 0$. But then $0 < -x = |x| = 0$, a contradiction. Hence, x must be equal to 0.

For part 2, let $x, y \in \mathbb{R}$. If $x \geq 0$ and $y \geq 0$, then by part 3 of Theorem 1.10 $xy \geq 0$, and hence $|xy| = xy = |x||y|$. If $x \geq 0$ and $y < 0$, then by part 4 of Theorem 1.10 we infer $xy \leq 0$. Hence, $|xy| = -xy = x(-y) = |x||y|$. The case $x < 0$ and $y \geq 0$ is similar and the reader will produce it in Exercise 1-11a. Finally, if $x < 0$ and $y < 0$, then by part 4 of Theorem 1.10 we obtain $xy > 0$. Hence, $|xy| = xy = (-1)(-1)xy = (-x)(-y) = |x||y|$.

To prove the triangular inequality, first note that for all $x \in \mathbb{R}$ we have that $x \leq |x|$. This is clear for $x \geq 0$ and for $x < 0$ we simply note $x < 0 < -x = |x|$. Moreover, (see Exercise 1-11b) for all $x \in \mathbb{R}$ we have $-x \leq |x|$. Now let $x, y \in \mathbb{R}$. If the inequality $x + y \geq 0$ holds, then by part 2 of Theorem 1.10 at least one of x, y is greater than or equal to 0. (Otherwise $x < 0$ and $y < 0$ would imply $x + y < 0$.) Hence, by part 2 of Theorem 1.10 $|x + y| = x + y \leq |x| + y \leq |x| + |y|$. If $x + y < 0$, then at least one of x and y is less than 0. Hence, by part 2 of Theorem 1.10 we obtain $|x + y| = -(x + y) = -x + (-y) \leq |-x| + (-y) \leq |-x| + |-y| = |x| + |y|$.

Finally, for the reverse triangular inequality, let $x, y \in \mathbb{R}$. Without loss of generality *(see Standard Proof Technique 1.14)* assume that $|x| \geq |y|$. (The proof for the case $|x| < |y|$ is left as Exercise 1-11c.) Then $|x| = |x - y + y| \leq |x - y| + |y|$, which implies $\big||x| - |y|\big| = |x| - |y| \leq |x - y|$. ■

Standard Proof Technique 1.14 If the proofs for the cases in a case distinction are very similar, it is customary to assume **without loss of generality** that one of these similar cases is true. This is not a loss of generality, because it is assumed that what is presented enables the reader to fill in the proof(s) for the other case(s). In this text, the omitted part is sometimes included as an explicit exercise for the reader. □

Exercises

1-10. Finishing the proof of Theorem 1.10.

(a) Prove part 1 of Theorem 1.10.

(b) Prove part 2 of Theorem 1.10.

1-11. Finishing the proof of Theorem 1.13.

(a) Let $x, y \in \mathbb{R}$. Prove that if $x \geq 0$ and $y < 0$, then $|xy| = |x||y|$.

(b) Prove that for all $x \in \mathbb{R}$ we have $-x \leq |x|$.

(c) Prove that if $|x| < |y|$, then $\big| |x| - |y| \big| \leq |x - y|$.

1-12. Let $I, J \subseteq \mathbb{R}$ be intervals. Prove that $I \cap J = \{x \in \mathbb{R} : x \in I \text{ and } x \in J\}$ is again an interval.

1-13. Let $a < b$ and let $x, y \in [a, b]$. Prove that $|x - y| \leq b - a$.

1-14. Prove that none of the fields from Exercise 1-9c can satisfy Axiom 1.6 by showing that for these fields part 2 of Axiom 1.6 fails for $x = 1$.

Note. This result shows that Axiom 1.6 distinguishes $\mathbb{R}$ from the finite fields of Exercise 1-9c.

1.3 Lowest Upper and Greatest Lower Bounds

A structure that has the properties outlined in Axioms 1.1 and 1.6 is also called a **linearly ordered field**. The rational numbers satisfy these properties just as well as the real numbers. Thus we are not done with our characterization of $\mathbb{R}$. The final axiom for the real numbers addresses upper and lower bounds of sets.

Definition 1.15 *Let A be a subset of $\mathbb{R}$.*

1. *The number $u \in \mathbb{R}$ is called an* **upper bound** *of A iff $u \geq a$ for all $a \in A$. If A has an upper bound, it is also called* **bounded above**.

2. *The number $l \in \mathbb{R}$ is called a* **lower bound** *of A iff $l \leq a$ for all $a \in A$. If A has a lower bound, it is also called* **bounded below**.

A subset $A \subseteq \mathbb{R}$ that is bounded above and bounded below is also called **bounded**.

Among all upper bounds of a set, the smallest one (if it exists) plays a special role. Similarly, the greatest lower bound plays a special role if it exists.

Definition 1.16 *Let $A \subseteq \mathbb{R}$.*

1. *The number $s \in \mathbb{R}$ is called* **lowest upper bound** *of A or* **supremum** *of A, denoted* $\sup(A)$, *iff s is an upper bound of A and for all upper bounds u of A we have that $s \leq u$.*

2. *The number $i \in \mathbb{R}$ is called* **greatest lower bound** *of A or* **infimum** *of A, denoted* $\inf(A)$, *iff i is a lower bound of A and for all lower bounds l of A we have that $l \leq i$.*

Formally, it is not guaranteed that suprema and infima are unique, but the next result shows that this is indeed the case. Note that the statement of Proposition 1.17 follows the standard pattern for a uniqueness statement.

Proposition 1.17 Suprema are unique. *That is, if the set $A \subseteq \mathbb{R}$ is bounded above and $s, t \in \mathbb{R}$ both are suprema of A, then $s = t$.*

Proof. Let $A \subseteq \mathbb{R}$ and $s, t \in \mathbb{R}$ be as indicated. Then s is an upper bound of A and, because t is a supremum of A, we infer $s \geq t$. Similarly, t is an upper bound of A and, because s is a supremum of A, we infer $t \geq s$. This implies $s = t$. ■

Standard Proof Technique 1.18 (Also compare with Standard Proof Technique 1.14.) When, as in the proof of Proposition 1.17, two parts of a proof are very similar, it is common to only prove one part and state that the other part is similar. Throughout the text, the reader will become familiar with this idea through exercises that require the construction of proofs that are similar to proofs given in the narrative. □

The proof that infima are unique is similar (see Exercise 1-15). Because suprema and infima are unique if they exist, we speak of *the* supremum and *the* infimum.

The final axiom for the real numbers now states that suprema and infima exist under mild hypotheses.

Axiom 1.19 Completeness Axiom. *Every nonempty subset S of $\mathbb{R}$ that has an upper bound has a lowest upper bound.*

Although the Completeness Axiom formally only guarantees that nonempty subsets of $\mathbb{R}$ that are bounded above have suprema, existence of infima is a consequence.

Proposition 1.20 *Let $S \subseteq \mathbb{R}$ be nonempty and bounded below. Then S has a greatest lower bound.*

Proof. Let $L := \{x \in \mathbb{R} : x \text{ is a lower bound of } S\}$. Then $L \neq \emptyset$. Let $s \in S$. Then for all $l \in L$ we have that $l \leq s$. Because $S \neq \emptyset$ this means that L is bounded above. Because $L \neq \emptyset$, by the Completeness Axiom, L has a supremum $\sup(L)$. Every $s \in S$ is an upper bound of L, which means that $s \geq \sup(L)$ and so $\sup(L)$ is a lower bound of S. By definition of suprema, $\sup(L)$ is greater than or equal to all elements of L, that is, it is greater than or equal to all lower bounds of S. By definition of infima, this means that $\sup(L) = \inf(S)$. ■

We will see that suprema and infima are valuable tools in analysis on the real line. The next result shows that in any set with a supremum we can find numbers that are arbitrarily close to the supremum. This fact is important, because analysis ultimately is about objects "getting close to each other."

Proposition 1.21 *Let $S \subset \mathbb{R}$ be a nonempty subset of $\mathbb{R}$ that is bounded above and let $s := \sup(S)$. Then for every $\varepsilon > 0$ there is an element $x \in S$ so that $s - x < \varepsilon$.*

Proof. Suppose for a contradiction that there is an $\varepsilon > 0$ so that for all $x \in S$ we have that $s - x \geq \varepsilon$. Then for all $x \in S$ we would obtain $s - \varepsilon \geq x$, that is, $s - \varepsilon$ would be an upper bound of S. But $s - \varepsilon < s$ contradicts the fact that s is the lowest upper bound of S. ■

Although the supremum and infimum of a set need not be elements of the set, we have different names for them in case they are in the set.

Definition 1.22 *Let A be a subset of $\mathbb{R}$.*

1. *If A is bounded above and* $\sup(A) \in A$, *then the supremum of A is also called the* **maximum** *of A, denoted* $\max(A)$.
2. *If A is bounded below and* $\inf(A) \in A$, *then the infimum of A is also called the* **minimum** *of A, denoted* $\min(A)$.

Although the distinctions between suprema and maxima and between infima and minima are small, the notions are distinct. For example, the open interval $(0, 1)$ has a supremum (1) and an infimum (0), but it has neither a maximum, nor a minimum.

Exercises

1-15. Let $A \subseteq \mathbb{R}$ be bounded below and let $s, t \in \mathbb{R}$ both be infima of A. Prove that $s = t$.

1-16. Approaching infima. State and prove a version of Proposition 1.21 that applies to infima. Is the proof significantly different from that of Proposition 1.21?

1-17. Let $S \subseteq \mathbb{R}$ be bounded above. Prove that $s \in \mathbb{R}$ is the supremum of S iff s is an upper bound of S and for all $\varepsilon > 0$ there is an $x \in S$ so that $|s - x| < \varepsilon$.

1-18. Suprema and infima vs. containment of sets.

(a) Let $A, B \subseteq \mathbb{R}$ be bounded above. Prove that $A \subseteq B$ implies $\sup(A) \leq \sup(B)$.

(b) Let $A, B \subseteq \mathbb{R}$ be bounded below. Prove that $A \subseteq B$ implies $\inf(A) \geq \inf(B)$.

1-19. Let $A \subseteq \mathbb{R}$ be bounded above. Prove that $\inf\{x \in \mathbb{R} : -x \in A\} = -\sup(A)$.

This section concludes the introduction of the axioms for the real numbers. Exercise 1-30 after the next section shows that the axioms *uniquely* determine the real numbers. We will not explicitly construct a mathematical entity that satisfies these axioms. Readers interested in the construction of $\mathbb{R}$ from $\mathbb{Q}$ can revisit this idea after Theorem 16.89 (see Exercise 16-93). The construction of the rational numbers from the axioms of set theory is sketched in Appendix C.

1.4 Natural Numbers, Integers, and Rational Numbers

Although Axioms 1.1, 1.6 and 1.19 uniquely describe the real numbers, they do not mention familiar subsets, such as natural numbers, integers, and rational numbers. This is because these sets can be constructed from the axioms as subsets of the real numbers. We start with the natural numbers, which are the unique subset with properties as stated in Theorem 1.23. While their existence is easy to establish, the uniqueness of the natural numbers can only be proved in Theorem 1.28 after some more machinery has been developed.

Theorem 1.23 *There is a subset $\mathbb{N} \subseteq \mathbb{R}$, called the* **natural numbers**, *so that*

1. $1 \in \mathbb{N}$.
2. *For each $n \in \mathbb{N}$ the number $n + 1$ is also in $\mathbb{N}$.*
3. **Principle of Induction**. *If $S \subseteq \mathbb{N}$ is such that $1 \in S$ and for each $n \in S$ we also have $n + 1 \in S$, then $S = \mathbb{N}$.*

Proof. Call a subset $A \subseteq \mathbb{R}$ a **successor set** iff $1 \in A$ and for all $a \in A$ we also have $a + 1 \in A$. Successor sets exist, because, for example, $\mathbb{R}$ itself is a successor set. Let $\mathbb{N}$ be the set of all elements of $\mathbb{R}$ that are in all successor sets. Because 1 is an element of every successor set, we infer $1 \in \mathbb{N}$. Moreover, if $n \in \mathbb{N}$, then n is in every successor set, which means $n + 1$ is in every successor set, and hence $n + 1 \in \mathbb{N}$. Finally, any subset $S \subseteq \mathbb{N}$ as given in the Principle of Induction is a successor set. Because the elements of $\mathbb{N}$ are contained in all successor sets, we conclude that $\mathbb{N} \subseteq S$, and hence $\mathbb{N} = S$. ■

Of course, we will denote the natural numbers by their usual names $1, 2, 3, \ldots$ As algebraic objects, natural numbers are suited for addition and multiplication (see Proposition 1.24), but they are not so well suited for subtraction (see Proposition 1.25). Although all results until Theorem 1.28 are stated for $\mathbb{N}$, they hold "for every subset of $\mathbb{R}$ that satisfies the properties in Theorem 1.23." The reader should keep this in mind and double check, because we will need it in the proof of Theorem 1.28. To avoid awkward formulations, the results up to Theorem 1.28 are formulated for $\mathbb{N}$, however.

Proposition 1.24 *The natural numbers are closed under addition and multiplication. That is, if $m, n \in \mathbb{N}$, then $m + n$ and mn are in $\mathbb{N}$ also.*

Proof. The key to this result is the Principle of Induction. Let $m \in \mathbb{N}$ be arbitrary and let $S_m := \{n \in \mathbb{N} : m+n \in \mathbb{N}\}$. Then $m \in \mathbb{N}$ implies $m+1 \in \mathbb{N}$, and hence $1 \in S_m$. Moreover, if $n \in S_m$, then $m + n \in \mathbb{N}$, and hence $m + (n + 1) = (m + n) + 1 \in \mathbb{N}$, which means that $n + 1 \in S_m$. By the Principle of Induction we conclude that $S_m = \mathbb{N}$. Because $m \in \mathbb{N}$ was arbitrary, this means that for any $m, n \in \mathbb{N}$ we have $m + n \in \mathbb{N}$.

The proof for products is similar and left to the reader as Exercise 1-20. ■

Readers familiar with induction recognize the part "$1 \in S_m$" of the preceding proof as the **base step** of an induction and the part "$n \in S_m \Rightarrow n + 1 \in S_m$" as the **induction step**. In this section, we use the "induction on sets" as done in the preceding proof. The more commonly known Principle of Induction is introduced in Theorem 1.39.

Proposition 1.25 *Let $m, n \in \mathbb{N}$ be such that $m > n$. Then $m - n \in \mathbb{N}$.*

Proof. We first show that if $m \in \mathbb{N}$, then $m - 1 \in \mathbb{N}$ or $m - 1 = 0$. To do this, let $A := \{m \in \mathbb{N} : m - 1 \in \mathbb{N} \text{ or } m - 1 = 0\}$. Then $1 \in A$ and if $m \in A$, then $(m + 1) - 1 = m \in A \subseteq \mathbb{N}$, which means $m + 1 \in A$. Hence, $A = \mathbb{N}$ by the Principle of Induction.

Now let $S := \{n \in \mathbb{N} : (\forall m \in \mathbb{N} : m > n \text{ implies } m - n \in \mathbb{N})\}$. If $n = 1$ and $m \in \mathbb{N}$ satisfies $m > 1$, then $m - 1 > 0$ and so by the above $m - 1 \in \mathbb{N}$, which means $1 \in S$. Let $n \in S$. If $m > n + 1$, then $m - 1 > n$, and hence $m - (n + 1) = (m - 1) - n \in \mathbb{N}$, which means $n + 1 \in S$. By the Principle of Induction we conclude that $S = \mathbb{N}$, and hence for all $m, n \in \mathbb{N}$ we have proved that $m > n$ implies $m - n \in \mathbb{N}$. ■

Proposition 1.26 shows that the natural numbers are positive and the smallest difference between any two of them is 1.

Proposition 1.26 *For all $n \in \mathbb{N}$, the inequality $n \geq 1$ holds and there is no $m \in \mathbb{N}$ so that the inequalities $n < m < n + 1$ hold.*

Proof. The proof that all natural numbers are greater than or equal to 1 is left to Exercise 1-21.

Now suppose for a contradiction that there is an $n \in \mathbb{N}$ and an $m \in \mathbb{N}$ so that $n < m < n + 1$. Then $m - n \in \mathbb{N}$ and $m - n < 1$, a contradiction. ■

The Well-ordering Theorem turns out to be equivalent to the Principle of Induction (see Exercise 1-22).

Theorem 1.27 Well-ordering Theorem. *Every nonempty subset of $\mathbb{N}$ has a smallest element.*

Proof. Suppose for a contradiction that $B \subseteq \mathbb{N}$ is not empty and does not have a smallest element. Let $S := \{n \in \mathbb{N} : (\forall m \in \mathbb{N} : m \leq n \text{ implies } m \notin B)\}$. By Proposition 1.26, 1 is less than or equal to all elements of $\mathbb{N}$, so $1 \notin B$, and hence $1 \in S$. Now let $n \in S$. Then all $m \in \mathbb{N}$ with $m \leq n$ are not in B. But then $n + 1 \in B$ would by Proposition 1.26 imply that $n + 1$ is the smallest element of B. Hence, $n + 1 \notin B$ and we conclude $n + 1 \in S$. By the Principle of Induction, $S = \mathbb{N}$ and consequently $B = \emptyset$, a contradiction. ■

Now we are finally ready to show that the natural numbers are unique.

Theorem 1.28 *The natural numbers $\mathbb{N}$ are the* unique *subset of $\mathbb{R}$ that satisfies the properties in Theorem 1.23.*

Proof. Examination of the proofs of all results since Theorem 1.23 reveals that any set $S \subseteq \mathbb{R}$ that satisfies the properties in Theorem 1.23 must also have the properties given in these results.

It may feel tedious to go back and verify the above statement. However, mathematical presentations more often than not will ask a reader to use a modification of a known proof to prove a result (also see Standard Proof Technique 1.14). When this occurs, the