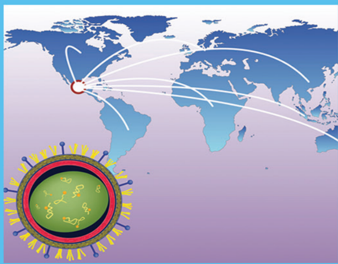
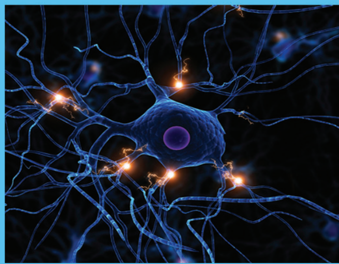
CHAPMAN & HALL/CRC APPLIED MATHEMATICS
AND NONLINEAR SCIENCE SERIES

Discovering Evolution Equations with Applications

Volume 2-Stochastic Equations





$$\begin{cases} du(t;\omega) = \left(Au(t;\omega) + f(u)(t;\omega)\right)dt + \\ \qquad\qquad g(u)(t;\omega)dW(t;\omega),\ (t;\omega) \in [0,T] \times \Omega, \\ u(0;\omega) = u_0(\omega), \omega \in \Omega \end{cases}$$

Mark A. McKibben

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Discovering Evolution Equations with Applications

Volume 2-Stochastic Equations

CHAPMAN & HALL/CRC APPLIED MATHEMATICS
AND NONLINEAR SCIENCE SERIES
Series Editor Goong Chen

Published Titles

Advanced Differential Quadrature Methods, Zhi Zong and Yingyan Zhang
Computing with hp-ADAPTIVE FINITE ELEMENTS, Volume 1, One and Two Dimensional Elliptic and Maxwell Problems, Leszek Demkowicz
Computing with hp-ADAPTIVE FINITE ELEMENTS, Volume 2, Frontiers: Three Dimensional Elliptic and Maxwell Problems with Applications, Leszek Demkowicz, Jason Kurtz, David Pardo, Maciej Paszyński, Waldemar Rachowicz, and Adam Zdunek
CRC Standard Curves and Surfaces with Mathematica®*: Second Edition,* David H. von Seggern
Discovering Evolution Equations with Applications: Volume 1-Deterministic Equations, Mark A. McKibben
Discovering Evolution Equations with Applications: Volume 2-Stochastic Equations, Mark A. McKibben
Exact Solutions and Invariant Subspaces of Nonlinear Partial Differential Equations in Mechanics and Physics, Victor A. Galaktionov and Sergey R. Svirshchevskii
Fourier Series in Several Variables with Applications to Partial Differential Equations, Victor L. Shapiro
Geometric Sturmian Theory of Nonlinear Parabolic Equations and Applications, Victor A. Galaktionov
Green's Functions and Linear Differential Equations: Theory, Applications, and Computation, Prem K. Kythe
Introduction to Fuzzy Systems, Guanrong Chen and Trung Tat Pham
Introduction to non-Kerr Law Optical Solitons, Anjan Biswas and Swapan Konar
Introduction to Partial Differential Equations with MATLAB®, Matthew P. Coleman
Introduction to Quantum Control and Dynamics, Domenico D'Alessandro
Mathematical Methods in Physics and Engineering with Mathematica, Ferdinand F. Cap
Mathematical Theory of Quantum Computation, Goong Chen and Zijian Diao
Mathematics of Quantum Computation and Quantum Technology, Goong Chen, Louis Kauffman, and Samuel J. Lomonaco
Mixed Boundary Value Problems, Dean G. Duffy
Multi-Resolution Methods for Modeling and Control of Dynamical Systems, Puneet Singla and John L. Junkins
Optimal Estimation of Dynamic Systems, John L. Crassidis and John L. Junkins
Quantum Computing Devices: Principles, Designs, and Analysis, Goong Chen, David A. Church, Berthold-Georg Englert, Carsten Henkel, Bernd Rohwedder, Marlan O. Scully, and M. Suhail Zubairy
A Shock-Fitting Primer, Manuel D. Salas
Stochastic Partial Differential Equations, Pao-Liu Chow

CHAPMAN & HALL/CRC APPLIED MATHEMATICS
AND NONLINEAR SCIENCE SERIES

Discovering Evolution Equations with Applications

Volume 2-Stochastic Equations

Mark A. McKibben

Goucher College

Baltimore, Maryland

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20110429

International Standard Book Number-13: 978-1-4200-9212-7 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Dedicated to my mother, Pat.

This page intentionally left blank

Contents

This page intentionally left blank

Preface

The mathematical modeling of complex phenomena that evolve over time relies heavily on the analysis of a variety of systems of ordinary and partial differential equations. Such models are developed in very disparate areas of study, ranging from the physical and natural sciences and population ecology to economics, neural networks, and infectious disease epidemiology. Despite the eclectic nature of the fields in which these models are formulated, various groups of them share enough common characteristics that make it possible to study them within a unified theoretical framework. Such study is an area of functional analysis commonly referred to as the theory of evolution equations.

In the absence of "noise," the evolution equations are said to be deterministic. If noise is taken into account in these models, by way of perturbations of the operators involved or via a Wiener process, then the evolution equations become stochastic in nature. The development of the general theory is similar to the deterministic case, but a considerable amount of additional machinery is needed in order to rigorously handle the addition of noise, and questions regarding the nature of the solutions (which are now viewed as stochastic processes rather than deterministic mappings) need to be addressed due to the probabilistic nature of the equations.

One thread of development in this vast field is the study of evolution equations that can be written in an abstract form analogous to a system of finite-dimensional linear ordinary differential equations. The ability to represent the solution of such a finite-dimensional system by a variation of parameters formula involving the matrix exponential prompts one, by analogy, to identify the entity that plays the role of the matrix exponential in a more abstract setting. Depending on the class of equations, this entity can be interpreted as a linear C_0-semigroup, a nonlinear semigroup, a (co)sine family, etc. A general theory is then developed in each situation and applied, to the extent possible, to all models within its parlance.

The literature for the theory of evolution equations is massive. Numerous monographs and journal articles have been written, the total sum of which covers a practically insurmountable amount of ground. While there exist five-volume magna opi that provide excellent accounts of the big picture of aspects of the field (for instance, **[105, 107, 418]**), most books written on evolution equations tend to either provide a thorough treatment of a particular class of equations in tremendous depth for a beginner or focus on presenting an assimilation of materials devoted to a very particular timely research direction (see **[11, 37, 38, 46, 47, 65, 90, 108, 131, 132, 133, 149, 159, 174, 178, 192, 206, 250, 252, 253, 290, 300, 305, 328, 329, 341, 365, 375, 381, 396, 407, 419]**). The natural practice in such mathematics texts, given that they are written for readers trained in advanced mathematics, is to pay little attention to preliminary material or behind-the-scenes detail. Needless to say, initiating study in this field can be daunting for beginners. This begs the question, "How do newcomers obtain an overview of the field, in a reasonable amount of time, that prepares them to enter and initially navigate the research realm?" This is what prompted me to embark on writing the current volume. The purpose of this volume is to provide an engaging, accessible account of a rudimentary core of theoretical results that should be understood by anyone studying stochastic evolution equations in a way that gradually builds the reader's intuition. To accomplish this task, I have opted to write the book using a so-called discovery approach, the ultimate goal of which is to engage you, the reader, in the actual mathematical enterprise of studying stochastic evolution equations. Some characteristics of this approach that you will encounter in the text are mentioned below.

What are the "discovery approach" features of the text?

I have tried to extract the essence of my approaches to teaching this material to newcomers to the field and conducting my own research, and incorporate these features into the actual prose of the text. For one, I pose questions of all types throughout the development of the material, from verifying details and illustrating theorems with examples to posing (and proving) conjectures of actual results and analyzing broad strokes that occur within the development of the theory itself. At times, the writing takes the form of a conversation with you, by way of providing motivation for a definition, or setting the stage for the next step of a theoretical development, or prefacing an important theorem with a plain-English explanation of it. I sometimes pose rhetorical questions to you as a lead-in to a subsequent section of the text. The inclusion of such discussion facilitates "seeing the big picture" of a theoretical development that I have found naturally connects its various stages. You are not left guessing why certain results are being developed or why a certain path is being followed. As a result, the exposition in the text, at times, may lack the "polished style" of a mathematical monograph, and the language used will be colloquial English rather than the standard mathematical language that you would encounter in a journal article. But, this style has the benefit of encouraging you to not simply passively read the text, but rather work through it, which is essential to obtaining a meaningful grasp of the material.

I deliberately begin each chapter with a discussion of models, many of which are studied in several chapters and modified along the way to motivate the particular theory to be developed in a given chapter. The intent is to illustrate how taking into account natural additional complexity gives rise to more complicated initial-boundary value problems that, in turn, are formulated using more general abstract evolution equations. This connectivity among different fields and the centrality of the theory of evolution equations to their study are illustrated on the cover of the text.

The driving force of the discussion is the substantive collection of more than 500 questions and exercises dispersed throughout the text. I have inserted questions of all types directly into the development of the chapters with the intention of having you pause and either process what has just been presented or react to a rhetorical question posed. You might be asked to supply details in an argument, verify a definition or theorem using a particular example, create a counterexample to show why an extension of a theorem from one setting to another fails, or conjecture and prove a result of your own based on previous material, etc. The questions, in essence, constitute much of the behind-the-scenes detail that goes into actually formulating the theory. In the spirit of the conversational nature of the text, I have included a section entitled *Guidance for Selected Exercises* at the end of the first nine chapters that provides two layers of hints for selected exercises. Layer one, labeled as "A Small Nudge in a Right Direction" is intended to help you get started if you are stumped. The idea is that you will re-attempt the exercise using the hint. If you find this hint insufficient, the second layer of hints, labeled as "An Additional Thrust in a Right Direction" provides a more substantive suggestion as to how to proceed. In addition to this batch of exercises, you will encounter more questions or directives enclosed in parentheses throughout all parts of the text. The purpose of these less formal, yet equally important questions is to alert you to when details are being omitted or to call your attention to a specific portion of a proof to which I want you to pay close attention. You will likely view the occurrence of these questions to be, at times, disruptive. And, this is exactly the point of including them! The tendency is to gloss over details when working through material as technical as this, but doing so too often will create gaps in understanding. It is my hope that the inclusion of the combination of the two layers of hints for the formal exercises and this frequent questioning will reduce any reluctance you might have in working through the text.

Finally, most chapters conclude with a section in which some of the models used to motivate the chapter are revisited, but are now modified in order to account for an additional complexity. The impetus is to direct your thinking toward what awaits you in the next chapter. This short, but natural, section is meant to serve as a connective link between chapters.

For whom is this book accessible?

It is my hope that anybody possessing a basic familiarity with the real numbers and at least an exposure to the most elementary of differential equations, be it a student, engineer, scientist, or mathematician specializing in a different area, can work through this text to gain an initial understanding of stochastic evolution equations, how they are used in practice, and more than twenty different areas of study to which the theory applies. Indeed, while the level of the mathematics discussed in the text is conventionally viewed as a topic that a graduate student would encounter after studying stochastic and functional analysis, all of the underlying tools of stochastic and functional analysis necessary to intelligently work through the text are included, chapter by chapter as they arise. This, coupled with the conversational style in which the text is written, should make the material naturally accessible to a broad audience.

What material does this text cover, in broad strokes?

The present volume consists of ten chapters. The text opens with two substantive chapters devoted to creating basic real and stochastic analysis "toolboxes," the purpose of which is to arm you with the bare essentials of real and stochastic analysis needed to work through the rest of the book. If you are familiar with the topics in the chapter, I suggest you peruse the chapter to get a feel for the notation and terminology prior to moving on.

Chapter 3 is devoted to the development of the theory for homogenous one-dimensional stochastic ODEs, while Chapter 4 immediately extends this theory to systems of homogenous linear stochastic ordinary differential equations. These chapters act as a springboard into the development of its abstract counterpart in a more general separable Hilbert space. The discussion proceeds to the case of linear homogenous abstract stochastic evolution equations in Chapter 5, and subsequently in the next two chapters to the nonhomogenous and semi-linear cases. The case in which the forcing term is a functional (acting from one function space to another) is addressed in Chapter 8, followed by a discussion of Sobolev-type stochastic evolution equations in Chapter 9. These latter two chapters have been recent active research areas. Finally, the last chapter is devoted to a brief discussion of several different directions involving accessible topics of active research.

For each class of equations, a core of theoretical results concerning the following main topics is developed: the existence and uniqueness of solutions (in a variety of senses) under various growth assumptions, continuous dependence upon initial data and parameters, convergence results of various kinds, and elementary stability results (in a variety of senses).

A substantive collection of mathematical models arising in areas such as heat conduction, advection, fluid flow through fissured rocks, transverse vibrations in extensible beams, thermodynamics, population ecology, pharmacokinetics, spatial pattern formation, pheromone transport, neural networks, and infectious disease epidemiology are developed in stages throughout the text. In fact, the reason for studying the class of abstract equations of a given chapter is motivated by first considering modified versions of the model(s) discussed in the previous chapter, and subsequently formulating the batch of newly created initial-boundary value problems in the form of the abstract equation to be studied in that chapter.

In order to get the most out of this text, I strongly encourage you to read it alongside of volume 1 **[295]** and to make deliberate step-by-step comparisons of the theory in the deterministic and stochastic settings.

About the book cover

You might very well be wondering about the significance of the text cover. Would you believe that it embodies the main driving force behind the text? Indeed, the initial-value problem in the middle from which all arrows emanate serves as a theoretical central hub that mathematically binds the models depicted by the illustrations on the cover, to name just a few. Each of the eight pictures illustrates a scenario described by a mathematical model (involving partial differential equations) that is studied in this text. As you work through the text, you will discover that all of these models can be written abstractly in the form of the initial-value problem positioned in the middle of the cover. So, despite the disparate nature of the fields in which these models arise, they can all be treated under the same theoretical umbrella. This is the power of the abstract theory developed in this text.

Reading from left to right, and top to bottom, the fields depicted in the pictures are as follows: air pollution, infectious disease epidemiology, neural networks, chemical kinetics, combustion, population dynamics, spatial pattern formation, and soil mechanics.

Acknowledgments

Writing this text has been one of the most positive and fulfilling experiences of my professional career thus far. I found various stages of writing this book to be truly energizing and uplifting, while others required me to plumb the very depths of my patience and perseverance. I truly realize that a project like this would have never come to fruition without the constant support, encouragement, and good humor of many colleagues, students, friends, and family. While it is virtually impossible to acknowledge each and every individual who has, in some way, influenced me in a manner that helped either to steer me toward writing this book or to navigate murky (and at times very choppy) waters during the writing phase, I would like to acknowledge several at this time.

First and foremost, there are two notable women who have been indelible sources of encouragement, energy, and support for as long as they have been in my life. My wife, to whom the first volume is dedicated, prompted me for years to write this book. And, once I actually took her advice and began the process, she never abandoned my urgent and recurrent needs for technological help, editorial expertise, or idea-sounding, and has never begrudged me for the momentary lack of patience, uttering of angst-provoked witty remarks, or necessary "idea jotting" at 3 am. And my mother, to whom this second volume is dedicated, and who has supported all of my scholastic and professional endeavors for as long as I can remember. She relentlessly encourages me to temper hard work with balance; she always seems to know when to provide a good-humored story to lighten a stressed mood and when to politely remind me to "take a break!" This book would never have materialized if it were not for both of you; your unwavering support of my endeavors is a significant driving force!

Next, I am extremely lucky to have had two outstanding mentors during my college years. My first exposure to college-level mathematics, by way of analysis, took place nearly two decades ago in a one-on-one tutorial with Dr. David Keck. His unbridled enthusiasm for teaching and learning mathematics and well-timed witty humor have been infectious. Achieving the depth of his passion and honing my skills to mirror his innate ability to teach mathematics are goals to which I will continue to aspire for the duration of my academic career. And, my dissertation advisor, Dr. Sergiu Aizicovici, who took me under his wing as a graduate student and introduced me to various facets of the study of evolution equations and the world of mathematical research. His abilities to make the area come to life, to help a newcomer navigate the practically insurmountable literature with ease, and to tolerate and provide meaningful answers to even the most rudimentary of questions (which I admittedly asked quite often!) in a way that honed my intuition surrounding the subject matter are among the many reasons why I chose to pursue this area of research. You both have left indelible imprints on my development as a mathematician and educator.

Many people have been kind enough to provide honest feedback during various stages of this project. My colleagues Dr. Robert Lewand, Dr. Scott Sibley, Dr. Bernadette Tutinas, Dr. Cynthia Young, and Dr. Jill Zimmerman all provided valuable comments on portions of the prospectus that undoubtedly led to a stronger proposal. My colleagues Dr. Micah Webster and Dr. David Keck provided valuable feedback on their impression of the probability theory chapter, which led to improved transition to subsequent chapters of the text. My students Shana Lieberman, Jennifer Jordan, and Jordan Yoder endured various portions of the manuscript by way of independent study. I am proud to report that all three of them survived the experience (and seemed to enjoy it) and identified their fair share of errors in the early versions of the text! My colleague Dr. Tom Kelliher provided invaluable TEX help during early stages of the project; thank you for helping such a TEX neophyte! A special thanks to my wife, Jodi, for reading the prospectus and manuscript, and who will never let me live down the fact that she found a mathematical error in the prospectus.

I would like to thank the entire Taylor & Francis team. To my editor, Bob Stern, thank you for approaching me about writing this text and for patiently guiding me through the process from beginning to end. To my project coordinator, Jessica Vakili, who answered my production, stylistic, and marketing questions in a very helpful and timely manner. To my project editor, Karen Simon, for keeping the publication of this book on track. To Kevin Craig for designing the awesome cover of this book. And to Shashi Kumar, who helped me to overcome various LATEX issues throughout the typesetting process.

And, last but not least, I would like to thank you, the reader, for embarking on this journey with me through an amazingly rich field of mathematics. I hope your study is as fulfilling as mine has been thus far.

Mark A. McKibben

Chapter 1

A Basic Analysis Toolbox

Overview

The purpose of this chapter is to provide you with a succinct, hands-on introduction to elementary analysis that focuses on notation, main definitions and results, and the techniques with which you should be comfortable prior to working through this text. Additional topics will be introduced throughout the text whenever needed. Little is assumed beyond a working knowledge of the properties of real numbers, the "freshmen calculus," and a tolerance for mathematical rigor. Keep in mind that the presentation is not intended to be a complete exposition of real analysis. You are encouraged to refer to texts devoted to more comprehensive treatments of analysis (see **[17, 67, 196, 197, 234, 236, 250, 301, 353, 357, 372]**).

1.1 Some Basic Mathematical Shorthand

Symbolism is used heftily in mathematical exposition. Careful usage of some basic notation can streamline the verbiage. Some of the common symbols used are as follows.

Let P and Q be statements. (If the statement P changes depending on the value of some parameter x, we denote this dependence by writing $P(x)$.)
1.) The statement "not P," called the *negation of* P, is denoted by "$\neg P$."
2.) The statement "P or Q" is denoted by "$P \vee Q$," while the statement "P and Q" is denoted by "$P \wedge Q$."
3.) The statement "If P, then Q" is called an *implication*, and is denoted by "$P \Longrightarrow Q$" (read "P implies Q"). Here, P is called the *hypothesis* and Q is the *conclusion*.
4.) The statement "P if, and only if, Q" is denoted by "P iff Q" or "$P \Longleftrightarrow Q$." Precisely, this means "$(P \Longrightarrow Q) \wedge (Q \Longrightarrow P)$."
5.) The statement "$Q \Longrightarrow P$" is the *converse* of "$P \Longrightarrow Q$."
6.) The statement "$\neg Q \Longrightarrow \neg P$" is the *contrapositive* of "$P \Longrightarrow Q$." These two statements are equivalent.
7.) The symbol "$\exists$" is an existential quantifier and is read as "there exists" or "there is at least one."

8.) The symbol "$\forall$" is a universal quantifier and is read as "for every" or "for any."

Exercise 1.1.1. Let P, Q, R, and S be statements.
i.) Form the negation of "$P \wedge (Q \wedge R)$."
ii.) Form the negation of "$\exists\ x$ such that $P(x)$ holds."
iii.) Form the negation of "$\forall x$, $P(x)$ holds."
iv.) Form the contrapositive of "$(P \wedge Q) \Longrightarrow (\neg R \vee S)$."

Remark. Implication is a transitive relation in the sense that

$$((P \Longrightarrow Q) \wedge (Q \Longrightarrow R)) \Longrightarrow (P \Longrightarrow R).$$

For instance, a sequence of algebraic manipulations used to solve an equation is technically such a string of implications from which we conclude that the values of the variable obtained in the last step are the solutions of the original equation. Mathematical proofs are comprised of strings of implications, albeit of a somewhat more sophisticated nature.

1.2 Set Algebra

Informally, a *set* can be thought of as a collection of objects (e.g., real numbers, vectors, matrices, functions, other sets, etc.); the contents of a set are referred to as its *elements*. We usually label sets using uppercase letters and their elements by lowercase letters. Three sets that arise often and for whom specific notation will be reserved are

$$\begin{aligned}
\mathbb{N} &= \{1,2,3,...\} \\
\mathbb{Q} &= \text{the set of all rational numbers} \\
\mathbb{R} &= \text{the set of all real numbers}
\end{aligned}$$

If P is a certain property and A is the set of all objects having property P, we write $A = \{x : x \text{ has } P\}$ or $A = \{x | x \text{ has } P\}$. A set with no elements is *empty*, denoted by Ø.

If A is not empty and a is an element of A, we denote this fact by "$a \in A$." If a is not an element of A, a fact denoted by "$a \notin A$," where is it located? This prompts us to prescribe a universal set $\mathscr{U}$ that contains all possible objects of interest in our discussion. The following definition provides an algebra of sets.

Definition 1.2.1. Let A and B be sets.
i.) A is a *subset* of B, written $A \subset B$, whenever $x \in A \Longrightarrow x \in B$.
ii.) A *equals* B, written $A = B$, whenever $(A \subset B) \wedge (B \subset A)$.
iii.) The *complement of A relative to B*, written $B \setminus A$, is the set $\{x | x \in B \wedge x \notin A\}$.

Specifically, the complement relative to $\mathscr{U}$ is denoted by $\widetilde{A}$.
iv.) The *union of A and B* is the set $A \cup B = \{x | x \in A \vee x \in B\}$.
v.) The *intersection of A and B* is the set $A \cap B = \{x | x \in A \wedge x \in B\}$.
vi.) $A \times B = \{(a,b) | a \in A \wedge b \in B\}$.

Proving set equality requires that we show two implications. Use this fact when appropriate to complete the following exercises.

Exercise 1.2.1. Let A, B, and C be sets. Prove the following:
i.) $A \subset B$ iff $\widetilde{B} \subset \widetilde{A}$
ii.) $A = (A \cap B) \cup (A \setminus B)$
iii.) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
iv.) $\widetilde{(A \cap B)} = \widetilde{A} \cup \widetilde{B}$ and $\widetilde{(A \cup B)} = \widetilde{A} \cap \widetilde{B}$

Exercise 1.2.2. Explain how you would prove $A \neq B$.

Exercise 1.2.3. Formulate an extension of Def. 1.2.1(iv) through (vi) that works for any finite number of sets.

It is often necessary to consider the union or intersection of more than two sets, possibly infinitely many. So, we need a succinct notation for unions and intersections of an arbitrary number of sets. Let $\Gamma \neq \emptyset$. (We think of the members of Γ as labels.) Suppose to each $\gamma \in \Gamma$, we associate a set A_γ. The collection of all these sets, namely $\mathscr{A} = \left\{A_\gamma | \gamma \in \Gamma\right\}$, is a *family of sets indexed by* Γ. We define

$$\bigcup_{\gamma \in \Gamma} A_\gamma = \left\{x | \exists \gamma \in \Gamma \text{ such that } x \in A_\gamma\right\}, \tag{1.1}$$

$$\bigcap_{\gamma \in \Gamma} A_\gamma = \left\{x | \forall \gamma \in \Gamma, x \in A_\gamma\right\}. \tag{1.2}$$

If $\Gamma = \mathbb{N}$, we write $\bigcup_{n=1}^{\infty}$ and $\bigcap_{n=1}^{\infty}$ in place of $\bigcup_{\gamma \in \Gamma}$ and $\bigcap_{\gamma \in \Gamma}$, respectively.

Exercise 1.2.4. Let A be a set and $\left\{A_\gamma | \gamma \in \Gamma\right\}$ a family of sets indexed by Γ. Prove
i.) $A \cap \bigcup_{\gamma \in \Gamma} A_\gamma = \bigcup_{\gamma \in \Gamma} \left(A \cap A_\gamma\right)$ and $A \cup \bigcap_{\gamma \in \Gamma} A_\gamma = \bigcap_{\gamma \in \Gamma} \left(A \cup A_\gamma\right)$
ii.) $\left(\bigcup_{\gamma \in \Gamma} A_\gamma\right)^{\sim} = \bigcap_{\gamma \in \Gamma} \widetilde{A_\gamma}$ and $\left(\bigcap_{\gamma \in \Gamma} A_\gamma\right)^{\sim} = \bigcup_{\gamma \in \Gamma} \widetilde{A_\gamma}$
iii.) $A \times \bigcup_{\gamma \in \Gamma} A_\gamma = \bigcup_{\gamma \in \Gamma} \left(A \times A_\gamma\right)$ and $A \times \bigcap_{\gamma \in \Gamma} A_\gamma = \bigcap_{\gamma \in \Gamma} \left(A \times A_\gamma\right)$
iv.) $\bigcap_{\gamma \in \Gamma} A_\gamma \subset A_{\gamma_0} \subset \bigcup_{\gamma \in \Gamma} A_\gamma, \forall \gamma_0 \in \Gamma$.

1.3 Functions

The concept of a function is central to the study of mathematics.

Definition 1.3.1. Let A and B be sets.
i.) A subset $f \subset A \times B$ satisfying
  a.) $\forall x \in A, \exists y \in B$ such that $(x,y) \in f$,
  b.) $(x,y_1) \in f \wedge (x,y_2) \in f \implies y_1 = y_2$,
  is called a *function from A into B*. We say f is *B-valued,* denoted by $f : A \to B$.
ii.) The set A is called the *domain* of f, denoted $\mathrm{dom}(f)$.
iii.) The *range* of f, denoted by $\mathrm{rng}(f)$, is given by $\mathrm{rng}(f) = \{f(x) | x \in A\}$.

Remarks.
1. Notation: When defining a function using an explicit formula, say $y = f(x)$, the notation $x \mapsto f(x)$ is often used to denote the function. Also, we indicate the general dependence on a variable using a dot, say $f(\cdot)$. If the function depends on two independent variables, we distinguish between them by using a different number of dots for each, say $f(\cdot, \cdot\cdot)$.
2. The term *mapping* is used synonymously with the term *function*.
3. $\mathrm{rng}(f) \subset B$.

Exercise 1.3.1. Precisely define what it means for two functions f and g to be equal.

The following classification plays a role in determining if a function is invertible.

Definition 1.3.2. $f : A \to B$ is called
i.) *one-to-one* if $f(x_1) = f(x_2) \Longrightarrow x_1 = x_2, \forall x_1, x_2 \in A$;
ii.) *onto* whenever $\mathrm{rng}(f) = B$.

We sometimes wish to apply functions in succession in the following sense.

Definition 1.3.3. Suppose that $f : \mathrm{dom}(f) \to A$ and $g : \mathrm{dom}(g) \to B$ with $\mathrm{rng}(g) \subset \mathrm{dom}(f)$. The *composition of f with g*, denoted $f \circ g$, is the function $f \circ g : \mathrm{dom}(g) \to A$ defined by $(f \circ g)(x) = f(g(x))$.

Exercise 1.3.2. Show that, in general, $f \circ g \neq g \circ f$.

Exercise 1.3.3. Let $f : \mathrm{dom}(f) \to A$ and $g : \mathrm{dom}(g) \to B$ be such that $f \circ g$ is defined. Prove
i.) If f and g are onto, then $f \circ g$ is onto.
ii.) If f and g are one-to-one, then $f \circ g$ is one-to-one.

At times, we need to compute the functional values for all members of a subset of the domain, or perhaps determine the subset of the domain whose collection of functional values is a prescribed subset of the range. These notions are made precise below.

Definition 1.3.4. Let $f : A \to B$.
i.) For $X \subset A$, the *image of X under f* is the set $f(X) = \{f(x) | x \in X\}$.
ii.) For $Y \subset B$, the *pre-image of Y under f* is the set

$$f^{-1}(Y) = \{x \in \mathscr{A} \,|\, \exists y \in Y \,\text{such that}\, y = f(x)\}.$$

The following related properties are useful.

Proposition 1.3.5. *Suppose* $f : A \to B$ *is a function,* X, X_1, X_2, *and* X_γ, $\gamma \in \Gamma$, *are all subsets of* A *and* Y, Y_1, Y_2, *and* Y_γ, $\gamma \in \Gamma$, *are all subsets of* B. *Then,*
i.) **a.)** $X_1 \subset X_2 \Longrightarrow f(X_1) \subset f(X_2)$
b.) $Y_1 \subset Y_2 \Longrightarrow f^{-1}(Y_1) \subset f^{-1}(Y_2)$
ii.) **a.)** $f\left(\bigcup_{\gamma\in\Gamma} X_\gamma\right) = \bigcup_{\gamma\in\Gamma} f\left(X_\gamma\right)$
b.) $f^{-1}\left(\bigcup_{\gamma\in\Gamma} Y_\gamma\right) = \bigcup_{\gamma\in\Gamma} f^{-1}\left(Y_\gamma\right)$
iii.) **a.)** $f\left(\bigcap_{\gamma\in\Gamma} X_\gamma\right) \subset \bigcap_{\gamma\in\Gamma} f\left(X_\gamma\right)$
b.) $f^{-1}\left(\bigcap_{\gamma\in\Gamma} Y_\gamma\right) = \bigcap_{\gamma\in\Gamma} f^{-1}\left(Y_\gamma\right)$
iv.) **a.)** $X \subset f^{-1}(f(X))$
b.) $f\left(f^{-1}(Y)\right) \subset Y$

Exercise 1.3.4.
i.) Prove Prop. 1.3.5.
ii.) Impose conditions on f that would yield equality in Prop. 1.3.5(iv)(a) and (b).

We often consider functions whose domains and ranges are subsets of $\mathbb{R}$. For such functions, the notion of monotonicity is often a useful characterization.

Definition 1.3.6. Let $f : \mathrm{dom}(f) \subset \mathbb{R} \to \mathbb{R}$ and suppose that $\varnothing \neq S \subset \mathrm{dom}(f)$. We say that f is
i.) *nondecreasing on* S whenever $x_1, x_2 \in S$ with $x_1 < x_2 \Longrightarrow f(x_1) \le f(x_2)$;
ii.) *nonincreasing on* S whenever $x_1, x_2 \in S$ with $x_1 < x_2 \Longrightarrow f(x_1) \ge f(x_2)$.

Remark. The prefix "non" in both parts of Def. 1.3.6 is removed when the inequality is strict.

The arithmetic operations of real-valued functions are defined in the natural way. For such functions, consider the following exercise.

Exercise 1.3.5. Suppose that $f : \mathrm{dom}(f) \subset \mathbb{R} \to \mathbb{R}$ and $g : \mathrm{dom}(g) \subset \mathbb{R} \to \mathbb{R}$ are nondecreasing (resp. nonincreasing) functions on their domains.
i.) Which of the functions $f+g$, $f-g$, $f \cdot g$, and $\frac{f}{g}$, if any, are nondecreasing (resp. nonincreasing) on their domains?
ii.) Assuming that $f \circ g$ is defined, must it be nondecreasing (resp. nonincreasing) on its domain?

1.4 The Space $(\mathbb{R}, |\cdot|)$

1.4.1 Order Properties

The basic arithmetic and order features of the real number system are likely familiar, even if you have not worked through its formal construction. For our purposes, we shall begin with a set $\mathbb{R}$ equipped with two operations, addition and multiplication, satisfying these algebraic properties:

(i) addition and multiplication are both commutative and associative,
(ii) multiplication distributes over addition,
(iii) adding zero to any real number yields the same real number,
(iv) multiplying a real number by 1 yields the same real number,
(v) every real number has a unique additive inverse, and
(vi) every nonzero real number has a unique multiplicative inverse.

Moreover, $\mathbb{R}$ equipped with the natural "<" ordering is an ordered field and obeys the following properties.

Proposition 1.4.1. (Order Features of $\mathbb{R}$)
For all $x,y,z \in \mathbb{R}$, the following are true:
i.) *Exactly one of the relationships $x = y$, $x < y$, or $y < x$ holds;*
ii.) $x < y \implies x + z < y + z$;
iii.) $(x < y) \wedge (y < z) \implies x < z$;
iv.) $(x < y) \wedge (c > 0) \implies cx < cy$;
v.) $(x < y) \wedge (c < 0) \implies cx > cy$;
vi.) $(0 < x < y) \wedge (0 < w < z) \implies 0 < xw < yz$.

The following is an immediate consequence of these properties and is often the underlying principle used when verifying an inequality.

Proposition 1.4.2. *If $x,y \in \mathbb{R}$ are such that $x < y + \varepsilon$, $\forall \varepsilon > 0$, then $x \le y$.*

Proof. Suppose not; that is, $y < x$. Observe that for $\varepsilon = \frac{x-y}{2} > 0$, $y + \varepsilon = \frac{x+y}{2} < x$. (Why?) This is a contradiction. Hence, it must be the case that $x \le y$. □

Remark. The above argument is a very simple example of a proof by contradiction. The strategy is to assume that the conclusion is false and then use this additional hypothesis to obtain an obviously false statement or a contradiction of another hypothesis in the claim. More information about elementary proof techniques can be found in **[372]**.

Exercise 1.4.1.
i.) Let $x,y > 0$. Prove that $xy \le \frac{x^2+y^2}{2}$.
ii.) Show that if $0 < x < y$, then $x^n < y^n$, $\forall n \in \mathbb{N}$.

1.4.2 Absolute Value

The above is a heuristic description of the familiar algebraic structure of $\mathbb{R}$. When equipped with a distance-measuring artifice, a deeper topological structure of $\mathbb{R}$ can be defined and studied. This is done with the help of the absolute value function.

Definition 1.4.3. For any $x \in \mathbb{R}$, the *absolute value* of x, denoted $|x|$, is defined by

$$|x| = \begin{cases} x, x \geq 0, \\ -x, x < 0. \end{cases}$$

This can be viewed as a measurement of distance between real numbers within the context of a number line. For instance, the solution set of the equation "$|x-2| = 3$" is the set of real numbers x that are "3 units away from 2," namely $\{-1, 5\}$.

Exercise 1.4.2. Determine the solution set for the following equations:
i.) $|x-3| = 0$
ii.) $|x+6| = 2$.

Proposition 1.4.4. *These properties hold for all $x, y, z \in \mathbb{R}$ and $a \geq 0$:*
i.) $-|x| = \min\{-x, x\} \leq x \leq \max\{-x, x\} = |x|$
ii.) $|x| \geq 0, \forall x \in \mathbb{R}$
iii.) $|x| = 0$ *iff* $x = 0$
iv.) $\sqrt{x^2} = |x|$
v.) $|xy| = |x|\,|y|$
vi.) $|x| \leq a$ *iff* $-a \leq x \leq a$
vii.) $|x+y| \leq |x| + |y|$
viii.) $|x-y| \leq |x-z| + |z-y|$
ix.) $\big|\,|x| - |y|\,\big| \leq |x-y|$
x.) $|x-y| < \varepsilon, \forall \varepsilon > 0 \Longrightarrow x = y$

Exercise 1.4.3. Prove Prop. 1.4.4.

Exercise 1.4.4. Let $n \in \mathbb{N}$ and $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n \in \mathbb{R}$. Prove:
i.) (Cauchy-Schwarz) $\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)$
ii.) (Minkowski) $\left(\sum_{i=1}^{n} (x_i + y_i)^2\right)^{1/2} \leq \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} + \left(\sum_{i=1}^{n} y_i^2\right)^{1/2}$
iii.) $\left|\sum_{i=1}^{n} x_i\right|^M \leq \left(\sum_{i=1}^{n} |x_i|\right)^M \leq n^{M-1} \sum_{i=1}^{n} |x_i|^M, \forall M \in \mathbb{N}$

1.4.3 Completeness Property of $(\mathbb{R}, |\cdot|)$

It turns out that $\mathbb{R}$ has a fundamental and essential property referred to as *completeness*, without which the study of analysis could not proceed. We introduce some terminology needed to state certain fundamental properties of $\mathbb{R}$.

Definition 1.4.5. Let $\varnothing \neq S \subset \mathbb{R}$.
i.) S is *bounded above* if $\exists u \in \mathbb{R}$ such that $x \leq u, \forall x \in S$;
ii.) $u \in \mathbb{R}$ is an *upper bound* of S (ub(S)) if $x \leq u, \forall x \in S$;
iii.) $u_0 \in \mathbb{R}$ is the *maximum* of S (max(S)) if u_0 is an ub(S) and $u_0 \in S$;
iv.) $u_0 \in \mathbb{R}$ is the *supremum* of S (sup(S)) if u_0 is an ub(S) and $u_0 \leq u$, for any other $u =$ub(S).

The following analogous terms can be defined by reversing the inequality signs in Def. 1.4.5: *bounded below*, *lower bound* of S (lb(S)), *minimum* of S (min(S)), and *infimum* of S (inf(S)).

Exercise 1.4.5. Formulate precise definitions of the above terms.

Exercise 1.4.6. Let $\varnothing \neq S \subset \mathbb{R}$.
i.) How would you prove that sup(S)$= \infty$?
ii.) Repeat (i) for inf(S)$= -\infty$.

Definition 1.4.6. A set $\varnothing \neq S \subset \mathbb{R}$ is *bounded* if $\exists M > 0$ such that $|x| \leq M$, $\forall x \in S$.

It can be formally shown that $\mathbb{R}$ possesses the so-called *completeness property*. The importance of this concept in the present and more abstract settings cannot be overemphasized. We state it in the form of a theorem to highlight its importance. Consult **[17, 234]** for a proof.

Theorem 1.4.7. *If $\varnothing \neq S \subset \mathbb{R}$ is bounded above, then $\exists u \in \mathbb{R}$ such that $u =$*sup*(S). We say $\mathbb{R}$ is complete.*

Remark. The duality between the statements concerning sup and inf leads to the formulation of the following alternate statement of the completeness property:

$$If\, \varnothing \neq T \subset \mathbb{R}\text{ }is\text{ }bounded\text{ }below, then\, \exists\, v \in \mathbb{R}\text{ }such\text{ }that\text{ }v = \inf(S). \tag{1.3}$$

Exercise 1.4.7. Prove that (1.3) is equivalent to Thrm 1.4.7.

Proposition 1.4.8. (Properties of inf and sup) *Let $\varnothing \neq S, T \subset \mathbb{R}$.*
i.) *Assume $\exists$ sup(S). Then, $\forall \varepsilon > 0$, $\exists x \in S$ such that* $\sup(S) - \varepsilon < x \leq \sup(S)$.
ii.) *If $S \subset T$ and $\exists \sup(T)$, then $\exists \sup(S)$ and* $\sup(S) \leq \sup(T)$.
iii.) *Let $S + T = \{s + t | s \in S \wedge t \in T\}$. If S and T are bounded above, then $\exists \sup(S + T)$ and it equals* $\sup(S) + \sup(T)$.
iv.) *Let $c \in \mathbb{R}$ and define $cS = \{cs | s \in S\}$. If S is bounded, then $\exists \sup(cS)$ given by*

$$\sup(cS) = \begin{cases} c \cdot \sup(S), & if\, c \geq 0, \\ c \cdot \inf(S), & if\, c < 0. \end{cases} \tag{1.4}$$

v.) *Let $\varnothing \neq S, T \subset (0, \infty)$ and define $S \cdot T = \{s \cdot t | s \in S \wedge t \in T\}$. If S and T are bounded above, then $\exists \sup(S \cdot T)$ and it equals* $\sup(S) \cdot \sup(T)$.

Proof. We prove (iii) and leave the others for you to verify as an exercise.
Because S and T are nonempty, $S + T \neq \emptyset$. Further, because

$$s + t \leq \sup(S) + \sup(T), \forall s \in S, t \in T, \tag{1.5}$$

it follows that $\sup(S) + \sup(T)$ is an upper bound of $(S + T)$. (Why?) Hence, $\exists \sup(S + T)$ and

$$\sup(S + T) \leq \sup(S) + \sup(T). \tag{1.6}$$

To establish the reverse inequality, let $\varepsilon > 0$. By Prop. 1.4.8, $\exists s_0 \in S$ and $t_0 \in T$ such that

$$\sup(S) - \frac{\varepsilon}{2} < s_0 \text{ and } \sup(T) - \frac{\varepsilon}{2} < t_0. \tag{1.7}$$

Consequently,

$$\sup(S) + \sup(T) - \varepsilon < s_0 + t_0 \leq \sup(S+T). \tag{1.8}$$

Thus, we conclude from Prop. 1.4.2 that

$$\sup(S) + \sup(T) \leq \sup(S+T). \tag{1.9}$$

Claim (iii) now follows from (1.6) and (1.9). (Why?) □

Exercise 1.4.8.
i.) Prove the remaining parts of Prop. 1.4.8.
ii.) Formulate statements analogous to those in Prop. 1.4.8 for infs. Indicate the changes that must be implemented in the proofs.

Remark. Prop 1.4.8(i) indicates that we can get "arbitrarily close" to $\sup(S)$ with elements of S. This is especially useful in convergence arguments.

1.4.4 Topology of $\mathbb{R}$

You have worked with open and closed intervals in calculus, but what do the terms *open* and *closed* mean? Is there any significant difference between them? The notion of an open set is central to the construction of a so-called *topology* on $\mathbb{R}$. Interestingly, many of the theorems from calculus are formulated on closed, bounded intervals for very good reason. As we proceed with our analysis of $\mathbb{R}$, you will see that many of these results are consequences of some fairly deep topological properties of $\mathbb{R}$ which, in turn, follow from the completeness property.

Definition 1.4.9. Let $S \subset \mathbb{R}$.
i.) x is an *interior point* (int pt) of S if $\exists \varepsilon > 0$ such that $(x-\varepsilon, x+\varepsilon) \subset S$.
ii.) x is a *limit point* (lim pt) of S if $\forall \varepsilon > 0$, $(x-\varepsilon, x+\varepsilon) \cap S$ is infinite.
iii.) x is a *boundary point* (bdry pt) of S if

$$\forall \varepsilon > 0, (x-\varepsilon, x+\varepsilon) \cap S \neq \emptyset \text{ and } (x-\varepsilon, x+\varepsilon) \cap \widetilde{S} \neq \emptyset.$$

iv.) The *boundary* of S is the set $\partial S = \{x \in \mathbb{R} | x \text{ is a bdry pt of } S\}$.
v.) The *interior* of S is the set $\text{int}(S) = \{x \in \mathbb{R} | x \text{ is an int pt of } S\}$.
vi.) The *derived set* of S is the set $S' = \{x \in \mathbb{R} | x \text{ is a lim pt of } S\}$.
vii.) The *closure* of S is the set $\text{cl}_{\mathbb{R}}(S) = S \cup S'$.
viii.) S is *open* if every point of S is an int pt of S.
ix.) S is *closed* if S contains all of its lim pts.

Illustrating these concepts using a number line can facilitate your understanding of them. Do so when completing the following exercise.

Exercise 1.4.9. For each of these sets S, compute int(S), S', and $\mathrm{cl}_{\mathbb{R}}(S)$. Also, determine if S is open, closed, both, or neither.
i.) $[1,5]$
ii.) $\mathbb{Q}$
iii.) $\left\{\frac{1}{n}|n \in \mathbb{N}\right\}$
iv.) $\mathbb{R}$
v.) $\emptyset$

It is not difficult to establish the following duality between a set and its complement. It is often a useful tool when proving statements about open and closed sets.

Proposition 1.4.10. *Let $S \subset \mathbb{R}$. S is open iff $\widetilde{S}$ is closed.*

Exercise 1.4.10. Verify the following properties of open and closed sets.
i.) Let $n \in \mathbb{N}$. If $G_1, \ldots, G_n$ is a finite collection of open sets, then $\bigcap_{k=1}^{n} G_k$ is open.
ii.) Let $n \in \mathbb{N}$. If $F_1, \ldots, F_n$ is a finite collection of closed sets, then $\bigcup_{k=1}^{n} F_k$ is closed.
iii.) Let $\Gamma \neq \emptyset$. If G_γ is open, $\forall \gamma \in \Gamma$, then $\bigcup_{\gamma\in\Gamma} G_\gamma$ is open.
iv.) Let $\Gamma \neq \emptyset$. If F_γ is closed, $\forall \gamma \in \Gamma$, then $\bigcap_{\gamma\in\Gamma} F_\gamma$ is closed.
v.) If $S \subset T$, then $\mathrm{int}(S) \subset \mathrm{int}(T)$.
vi.) If $S \subset T$, then $\mathrm{cl}_{\mathbb{R}}(S) \subset \mathrm{cl}_{\mathbb{R}}(T)$.

Exercise 1.4.11. Let $\emptyset \neq S \subset \mathbb{R}$. Prove the following:
i.) If S is bounded above, then $\sup(S) \in \mathrm{cl}_{\mathbb{R}}(S)$.
ii.) If S is bounded above and closed, then $\max(S) \in S$.
iii.) Formulate results analogous to (i) and (ii) assuming that S is bounded below.

Intuitively, S' is the set of points to which elements of S become arbitarily close. It is natural to ask if there are proper subsets of $\mathbb{R}$ that sprawl widely enough through $\mathbb{R}$ as to be sufficiently near every real number. Precisely, consider sets of the following type.

Definition 1.4.11. A set $\emptyset \neq S \subset \mathbb{R}$ is *dense* in $\mathbb{R}$ if $\mathrm{cl}_{\mathbb{R}}(S) = \mathbb{R}$.

Exercise 1.4.12. Identify two different subsets of $\mathbb{R}$ that are dense in $\mathbb{R}$.

By way of motivation for the first major consequence of completeness, consider the following exercise.

Exercise 1.4.13. Provide examples, if possible, of sets $S \subset \mathbb{R}$ illustrating the following scenarios.
i.) S is bounded, but $S' = \emptyset$.
ii.) S is infinite, but $S' = \emptyset$.
iii.) S is bounded and infinite, but $S' = \emptyset$.

As you discovered in Exercise 1.4.13, the combination of bounded and infinite for a set S of real numbers implies the existence of a limit point of S. This is a consequence of the following theorem due to Bolzano and Weierstrass.

Theorem 1.4.12. (Bolzano-Weierstrass) *If S is a bounded, infinite subset of $\mathbb{R}$, then $S' \neq \emptyset$.*

Outline of Proof: Let $T = \{x \in \mathbb{R} | S \cap (x, \infty) \text{ is infinite}\}$. Then,

$$T \neq \emptyset. \text{ (Why?)} \tag{1.10}$$
$$T \text{ is bounded above. (Why?)} \tag{1.11}$$
$$\exists \sup(T); \text{ call it } t. \text{ (Why?)} \tag{1.12}$$
$$\forall \varepsilon > 0, S \cap (t - \varepsilon, \infty) \text{ is infinite. (Why?)} \tag{1.13}$$
$$\forall \varepsilon > 0, S \cap [t + \varepsilon, \infty) \text{ is finite. (Why?)} \tag{1.14}$$
$$\forall \varepsilon > 0, S \cap (t - \varepsilon, t + \varepsilon) \text{ is infinite. (Why?)} \tag{1.15}$$
$$t \in S'. \text{ (Why?)} \tag{1.16}$$

This completes the proof. □

Exercise 1.4.14. Provide the details in the proof of Thrm. 1.4.12. Where was completeness used?

Another important concept is that of compactness. Some authors define this notion more generally using open covers (see **[17]**).

Definition 1.4.13. A set $S \subset \mathbb{R}$ is *compact* if every infinite subset of S has a limit point <u>in S</u>.

Remark. The "in S" portion of Def. 1.4.13 is crucial, and it distinguishes between the sets $(0,1)$ and $[0,1]$, for instance. (Why?) This is evident in Thrm. 1.4.14.

Exercise 1.4.15. Try to determine if the following subsets of $\mathbb{R}$ are compact.
i.) Any finite set.
ii.) $\left\{\frac{1}{n} | n \in \mathbb{N}\right\}$ versus $\left\{\frac{1}{n} | n \in \mathbb{N}\right\} \cup \{0\}$
iii.) $\mathbb{Q}$
iv.) $\mathbb{Q} \cap [0,1]$
v.) $\mathbb{N}$
vi.) $\mathbb{R}$
vii.) $(0,1)$ versus $[0,1]$

Both the completeness property and finite dimensionality of $\mathbb{R}$ enter into the proof of the following characterization theorem for compact subsets of $\mathbb{R}$. The proof can be found in **[17]**.

Theorem 1.4.14. (Heine-Borel) *A set $S \subset \mathbb{R}$ is compact iff S is closed and bounded.*

Exercise 1.4.16. Revisit Exer. 1.4.15 in light of Thrm. 1.4.14.

1.5 Sequences in $(\mathbb{R}, |\cdot|)$

Sequences play a prominent role in analysis, especially in the development of numerical schemes used for approximation purposes.

1.5.1 Sequences and Subsequences

Definition 1.5.1. A *sequence* in $\mathbb{R}$ is a function $x : \mathbb{N} \to \mathbb{R}$. We often write x_n for $x(n)$, $n \in \mathbb{N}$, called the n^{th}-*term* of the sequence, and denote the sequence itself by $\{x_n\}$ or by enumerating the range as $x_1, x_2, x_3, \ldots$.

The notions of monotonicity and boundedness given in Defs. 1.3.6 and 1.4.6 apply in particular to sequences. We formulate them in this specific setting for later reference.

Definition 1.5.2. A sequence is called
i.) *nondecreasing* whenever $x_n \leq x_{n+1}$, $\forall n \in \mathbb{N}$;
ii.) *increasing* whenever $x_n < x_{n+1}$, $\forall n \in \mathbb{N}$;
iii.) *nonincreasing* whenever $x_n \geq x_{n+1}$, $\forall n \in \mathbb{N}$;
iv.) *decreasing* whenever $x_n > x_{n+1}$, $\forall n \in \mathbb{N}$;
v.) *monotone* if any of (i)–(iv) are satisfied;
vi.) *bounded above* (resp. *below*) if $\exists M \in \mathbb{R}$ such that $x_n \leq M$ (resp. $x_n \geq M$), $\forall n \in \mathbb{N}$;
vii.) *bounded* whenever $\exists M > 0$ such that $|x_n| \leq M$, $\forall n \in \mathbb{N}$.

Exercise 1.5.1. Explain why a nondecreasing (resp. nonincreasing) sequence must be bounded below (resp. above).

Definition 1.5.3. If $x : \mathbb{N} \to \mathbb{R}$ is a sequence in $\mathbb{R}$ and $n : \mathbb{N} \to \mathbb{N}$ is an increasing sequence in $\mathbb{N}$, then the composition $x \diamond n : \mathbb{N} \to \mathbb{R}$ is called a *subsequence* of x in $\mathbb{R}$.

Though this is a formal definition of a subsequence, let us examine carefully what this means using more conventional notation. Suppose that the terms of Def. 1.5.3 are represented by $\{x_n\}$ and $\{n_k\}$, respectively. Because $\{n_k\}$ is increasing, we know that $n_1 < n_2 < n_3 < \ldots$. Then, the official subsequence $x \circ n$ has values $(x \circ n)(k) = x(n(k))$, which, using our notation, can be written as x_{n_k}, $\forall k \in \mathbb{N}$. Thus, the integers n_k are just the indices of those terms of the original sequence that are retained in the subsequence as k increases, and roughly speaking, the remainder of the terms are omitted.

1.5.2 Limit Theorems

We now consider the important notion of convergence.

Definition 1.5.4. A sequence $\{x_n\}$ has *limit* L whenever $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ (N depending in general on ε) such that

$$n \geq N \implies |x_n - L| < \varepsilon.$$

In such case, we write $\lim_{n\to\infty} x_n = L$ or $x_n \longrightarrow L$ and say that $\{x_n\}$ *converges* (or *is convergent) to L*. Otherwise, we say $\{x_n\}$ *diverges*.

If we paraphrase Def. 1.5.4, it would read: $\lim_{n\to\infty} x_n = L$ whenever given <u>any</u> open interval $(L-\varepsilon, L+\varepsilon)$ around L (that is, no matter how small the positive number ε is), it is the case that $x_n \in (L-\varepsilon, L+\varepsilon)$ for all but possibly finitely many indices n. That is, the "tail" of the sequence ultimately gets into every open interval around L. Also note that, in general, the smaller the ε, the larger the index N must be used (to get deeper into the tail) because ε is an error gauge, namely how far the terms are from the target. We say N must be chosen "sufficiently large" as to ensure the tail behaves in this manner for the given ε.

Exercise 1.5.2.
i.) Precisely define $\lim_{n\to\infty} x_n \neq L$.
ii.) Prove that $x_n \longrightarrow L$ iff $|x_n - L| \longrightarrow 0$.

Example. As an illustration of Def. 1.5.4, we prove that $\lim_{n\to\infty} \frac{2n^2+n+5}{n^2+1} = 2$.
Let $\varepsilon > 0$. We must argue that $\exists N \in \mathbb{N}$ such that

$$n \geq N \Longrightarrow \left| \frac{2n^2+n+5}{n^2+1} - 2 \right| < \varepsilon. \tag{1.17}$$

To this end, note that $\exists N \in \mathbb{N}$ such that $N > 3$ and $N\varepsilon > 2$. (Why?) We show this N "works." Indeed, observe that $\forall n \geq N$,

$$\left| \frac{2n^2+n+5}{n^2+1} - 2 \right| = \left| \frac{2n^2+n+5-2n^2-2}{n^2+1} \right| = \frac{n+3}{n^2+1}. \tag{1.18}$$

Subsequently, by choice of N, we see that $n \geq N > 3$ and for all such n,

$$\frac{n+3}{n^2+1} < \frac{2n}{n^2+1} < \frac{2n}{n^2} = \frac{2}{n} < \frac{2}{N} < \varepsilon. \tag{1.19}$$

(Why?) Thus, by definition, it follows that $\lim_{n\to\infty} \frac{2n^2+n+5}{n^2+1} = 2$. □

Exercise 1.5.3. Use Def. 1.5.4 to prove that $\lim_{n\to\infty} \frac{a}{n} = 0$, $\forall a \in \mathbb{R}$.

We now discuss the main properties of convergence. We mainly provide outlines of proofs, the details of which you are encouraged to provide.

Proposition 1.5.5. *If $\{x_n\}$ is a convergent sequence, then its limit is unique.*

Outline of Proof: Let $\lim_{n\to\infty} x_n = L_1$ and $\lim_{n\to\infty} x_n = L_2$ and suppose that, by way of contradiction, $L_1 \neq L_2$.
Let $\varepsilon = \frac{|L_1-L_2|}{2}$. Then, $\varepsilon > 0$. (Why?)
$\exists N_1 \in \mathbb{N}$ such that $n \geq N_1 \Longrightarrow |x_n - L_1| < \varepsilon$. (Why?)

$\exists N_2 \in \mathbb{N}$ such that $n \geq N_2 \implies |x_n - L_2| < \varepsilon$. (Why?)
Choose $N = \max\{N_1, N_2\}$. Then, $|x_N - L_1| < \varepsilon$ and $|x_N - L_2| < \varepsilon$. (Why?)
Consequently, $2\varepsilon = |L_1 - L_2| \leq |x_N - L_1| + |x_N - L_2| < 2\varepsilon$. (Why?)
Thus, $L_1 = L_2$. (How?)
This completes the proof. □

Proposition 1.5.6. *If $\{x_n\}$ is a convergent sequence, then it is bounded.*

Outline of Proof: Assume that $\lim_{n\to\infty} x_n = L$. We must produce an $M > 0$ such that $|x_n| \leq M$, $\forall n \in \mathbb{N}$. Using $\varepsilon = 1$ in Def. 1.5.4, we know that

$$\exists N \in \mathbb{N} \text{ such that } n \geq N \implies |x_n - L| < \varepsilon = 1. \tag{1.20}$$

Using Prop. 1.4.4(ix) in (1.20) then yields

$$|x_n| < |L| + 1, \forall n \geq N. \tag{1.21}$$

(Tell how.) For how many values of n does x_n possibly not satisfy (1.21)? How do you use this fact to construct a positive real number M satisfying Def. 1.5.2(vii)? □

Proposition 1.5.7. (Squeeze Theorem) *Let $\{x_n\}, \{y_n\}$, and $\{z_n\}$ be sequences such that*

$$x_n \leq y_n \leq z_n, \forall n \in \mathbb{N}, \tag{1.22}$$

and

$$\lim_{n\to\infty} x_n = L = \lim_{n\to\infty} z_n. \tag{1.23}$$

Then, $\lim_{n\to\infty} y_n = L$.

Outline of Proof: Let $\varepsilon > 0$. From (1.23) we know that $\exists N_1, N_2 \in \mathbb{N}$ such that

$$|x_n - L| < \varepsilon, \forall n \geq N_1 \text{ and } |z_n - L| < \varepsilon, \forall n \geq N_2. \tag{1.24}$$

Specifically,

$$-\varepsilon < x_n - L, \forall n \geq N_1 \text{ and } z_n - L < \varepsilon, \forall n \geq N_2. \tag{1.25}$$

Choose $N = \max\{N_1, N_2\}$. Using (1.25) we see that

$$-\varepsilon < x_n - L, \text{ and } z_n - L < \varepsilon, \forall n \geq N. \text{ (Why?)}$$

Using this with (1.22) we can conclude that

$$n \geq N \implies -\varepsilon < y_n - L < \varepsilon. \text{ (Why?)}$$

Hence, $\lim_{n\to\infty} y_n = L$, as desired. □

Remark. The conclusion of Prop. 1.5.7 holds true if we replace (1.22) by

$$\exists N_0 \in \mathbb{N} \text{ such that } x_n \leq y_n \leq z_n, \forall n \geq N_0. \tag{1.26}$$

Suitably modify the way N is chosen in the proof of Prop. 1.5.7 to account for this more general condition. (Tell how.)

Proposition 1.5.8. *If* $\lim_{n\to\infty} x_n = L$, *where* $L \neq 0$, *then* $\exists m > 0$ *and* $N \in \mathbb{N}$ *such that*

$$|x_n| > m, \forall n \geq N.$$

(In words, if a sequence has a nonzero limit, then its terms must be bounded away from zero for sufficiently large indices n.)

Outline of Proof:

Let $\varepsilon = \frac{|L|}{2}$. Then, $\varepsilon > 0$. (Why?)

$\exists N \in \mathbb{N}$ such that $|x_n - L| < \varepsilon = \frac{|L|}{2}, \forall n \geq N$. (Why?)

Thus, $||x_n| - |L|| < \frac{|L|}{2}, \forall n \geq N$. (Why?)

That is, $-\frac{|L|}{2} < |x_n| - |L| < \frac{|L|}{2}, \forall n \geq N$. (Why?)

So, $\frac{|L|}{2} < |x_n|, \forall n \geq N$.

The conclusion follows by choosing $m = \frac{|L|}{2}$. (Why?) □

Proposition 1.5.9. *Suppose that* $\lim_{n\to\infty} x_n = L$ *and* $\lim_{n\to\infty} y_n = M$. *Then,*

i.) $\lim_{n\to\infty} (x_n + y_n) = L + M$;

ii.) $\lim_{n\to\infty} x_n y_n = LM$.

Outline of Proof:

Proof of (i): The strategy is straightforward. Because there are two sequences, we split the given error tolerance ε into two parts of size $\frac{\varepsilon}{2}$ each, apply the limit definition to each sequence with the $\frac{\varepsilon}{2}$ tolerance, and finally put the two together using the triangle inequality.

Let $\varepsilon > 0$. Then, $\frac{\varepsilon}{2} > 0$. We know that

$$\exists N_1 \in \mathbb{N} \text{ such that } |x_n - L| < \frac{\varepsilon}{2}, \forall n \geq N_1. \text{ (Why?)} \tag{1.27}$$

$$\exists N_2 \in \mathbb{N} \text{ such that } |y_n - M| < \frac{\varepsilon}{2}, \forall n \geq N_2. \text{ (Why?)} \tag{1.28}$$

How do you then select $N \in \mathbb{N}$ such that (1.27) and (1.28) hold simultaneously for all $n \geq N$? For such an N, observe that

$$n \geq N \implies |(x_n + y_n) - (L + M)| \leq |x_n - L| + |y_n - M| < \varepsilon. \tag{1.29}$$

(Why?) Hence, we conclude that $\lim_{n\to\infty} (x_n + y_n) = L + M$.

Proof of (ii): This time the strategy is a bit more involved. We need to show that $|x_n y_n - LM|$ can be made arbitrarily small for sufficiently large n using the hypotheses that $|x_n - L|$ and $|y_n - M|$ can each be made arbitrarily small for sufficiently large n. This requires two approximations, viz., making x_n close to L while simultaneously making y_n close to M. This suggests that we bound $|x_n y_n - LM|$ above by an expression involving $|x_n - L|$ and $|y_n - M|$. To accomplish this, we add and subtract the same middle term in $|x_n y_n - LM|$ and apply certain absolute value properties. Precisely, observe that

$$\begin{aligned} |x_n y_n - LM| &= |x_n y_n - M x_n + M x_n - LM| \\ &= |x_n (y_n - M) + M (x_n - L)| \\ &\leq |x_n| \, |y_n - M| + |M| \, |x_n - L| \, . \end{aligned} \tag{1.30}$$

(This trick is a workhorse throughout the text!) The tack now is to show that both terms on the right-hand side of (1.30) can be made less than $\frac{\varepsilon}{2}$ for sufficiently large n.

Let $\varepsilon > 0$. Proposition 1.5.6 implies that $\exists K > 0$ for which

$$|x_n| \leq K, \forall n \in \mathbb{N}. \tag{1.31}$$

Also, because $\{y_n\}$ is convergent to M, $\exists N_1 \in \mathbb{N}$ such that

$$|y_n - M| < \frac{\varepsilon}{2K}, \forall n \geq N_1. \tag{1.32}$$

In view of (1.31) and (1.32), we obtain

$$n \geq N_1 \implies |x_n| \, |y_n - M| \leq K \, |y_n - M| < K \cdot \frac{\varepsilon}{2K} = \frac{\varepsilon}{2}. \tag{1.33}$$

This takes care of the first term in (1.30). Next, because $\{x_n\}$ is convergent to L, $\exists N_2 \in \mathbb{N}$ such that

$$|x_n - L| < \frac{\varepsilon}{2\,(|M|+1)}, \forall n \geq N_2. \tag{1.34}$$

Now, argue in a manner similar to (1.33) to conclude that

$$n \geq N_2 \implies |M| \, |x_n - L| < \frac{\varepsilon}{2}. \text{ (Tell how.)} \tag{1.35}$$

Choose $N \in \mathbb{N}$ so that (1.33) and (1.35) hold simultaneously, $\forall n \geq N$. Then, use (1.30) through (1.35) to conclude that

$$n \geq N \implies |x_n y_n - LM| < \varepsilon. \text{ (How?)}$$

Hence, we conclude that $\lim_{n \to \infty} x_n y_n = LM$. This completes the proof. □

Because we were unfolding the argument in somewhat reverse order for motivation, it would be better now to start with $\varepsilon > 0$ and reorganize the train of the suggested argument into a polished proof. (Do so!)

Exercise 1.5.4. Let $c \in \mathbb{R}$ and assume that $\lim_{n \to \infty} x_n = L$ and $\lim_{n \to \infty} y_n = M$. Prove that

i.) $\lim_{n \to \infty} c x_n = cL$,

ii.) $\lim_{n \to \infty} (x_n - y_n) = L - M$.

The following lemma can be proven easily using induction. (Tell how.)

Lemma 1.5.10. *If $\{n_k\} \subset \mathbb{N}$ is an increasing sequence, then $n_k \geq k$, $\forall k \in \mathbb{N}$.*

Proposition 1.5.11. *If $\lim_{n\to\infty} x_n = L$ and $\left\{x_{n_k}\right\}$ is any subsequence of $\{x_n\}$, then $\lim_{k\to\infty} x_{n_k} = L$. (In words, all subsequences of a sequence convergent to L also converge to L.)*

Outline of Proof: Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that

$$|x_n - L| < \varepsilon, \forall n \geq N.$$

Now, fix any $K_0 \geq N$ and use Lemma 1.5.10 to infer that

$$k \geq K_0 \Longrightarrow n_k > k \geq K_0 \geq N \Longrightarrow \left|x_{n_k} - L\right| < \varepsilon. \text{ (Why?)}$$

The conclusion now follows. (Tell how.) □

Exercise 1.5.5. Prove that if $\lim_{n\to\infty} x_n = 0$ and $\{y_n\}$ is bounded, then $\lim_{n\to\infty} x_n y_n = 0$.

Exercise 1.5.6.
i.) Prove that if $\lim_{n\to\infty} x_n = L$, then $\lim_{n\to\infty} |x_n| = |L|$.
ii.) Provide an example of a sequence $\{x_n\}$ for which $\exists \lim_{n\to\infty} |x_n|$, but $\nexists \lim_{n\to\infty} x_n$.

Exercise 1.5.7. Prove the following:
i.) If $\lim_{n\to\infty} x_n = L$, then $\lim_{n\to\infty} x_n^p = L^p$, $\forall p \in \mathbb{N}$.
ii.) If $x_n > 0$, $\forall n \in \mathbb{N}$, and $\lim_{n\to\infty} x_n = L$, then $\lim_{n\to\infty} \sqrt{x_n} = \sqrt{L}$.

Important connections between sequences and the derived set and closure of a set are provided in the following exercise.

Exercise 1.5.8. Let $\varnothing \neq S \subset \mathbb{R}$. Prove the following:
i.) $x \in S'$ iff $\exists \{x_n\} \subset S \setminus \{x\}$ such that $\lim_{n\to\infty} x_n = x$.
ii.) $x \in \mathrm{cl}_{\mathbb{R}}(S)$ iff $\exists \{x_n\} \subset S$ such that $\lim_{n\to\infty} x_n = x$.

Proposition 1.5.12. *If $\{x_n\}$ is a bounded sequence in $\mathbb{R}$, then there exists a convergent subsequence $\left\{x_{n_k}\right\}$ of $\{x_n\}$.*

Outline of Proof: Let $\mathscr{R}_x = \{x_n | n \in \mathbb{N}\}$. We split the proof into two cases.

Case 1: $\mathscr{R}_x$ is a finite set, say $\mathscr{R}_x = \{y_1, y_2, \ldots, y_m\}$.

It cannot be the case that the set $x^{-1}(\{y_i\}) = \{n \in \mathbb{N} | x_n = y_i\}$ is finite, for every $i \in \{1,2,\ldots,m\}$ because $\mathbb{N} = \bigcup_{i=1}^m x^{-1}(\{y_i\})$. (Why?) As such, there is at least one $i_0 \in \{1,2,\ldots,m\}$ such that $x^{-1}\left(\left\{y_{i_0}\right\}\right)$ is infinite. Use this fact to inductively construct a sequence $n_1 < n_2 < \ldots$ in $\mathbb{N}$ such that $x_{n_k} = y_{i_0}$, $\forall k \in \mathbb{N}$. (Tell how.) Observe that $\left\{x_{n_k}\right\}$ is a convergent subsequence of $\{x_n\}$. (Why?)

Case 2: $\mathscr{R}_x$ is infinite.

Because $\{x_n\}$ is bounded, it follows from Thrm. 1.4.12 that $\mathscr{R}'_x \neq \varnothing$, say $L \in \mathscr{R}'_x$. Use the definition of limit point to inductively construct a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \longrightarrow L$. How does this complete the proof? □

The combination of the hypotheses of monotonicity and boundedness implies convergence, as the next result suggests.

Proposition 1.5.13. *If $\{x_n\}$ is a nondecreasing sequence that is bounded above, then $\{x_n\}$ converges and* $\lim_{n\to\infty} x_n = \sup\{x_n | n \in \mathbb{N}\}$.

Outline of Proof: Because $\{x_n | n \in \mathbb{N}\}$ is a nonempty subset of $\mathbb{R}$ that is bounded above, $\exists \sup\{x_n | n \in \mathbb{N}\}$, call it L. (Why?) Let $\varepsilon > 0$. Then,

$$\exists N \in \mathbb{N} \text{ such that } L - \varepsilon < x_N. \text{ (Why?)}$$

Consequently,

$$n \geq \mathbb{N} \Longrightarrow L - \varepsilon < x_N \leq x_n \leq L < L + \varepsilon \Longrightarrow |x_n - L| < \varepsilon.$$

(Why?) This completes the proof. □

Exercise 1.5.9. Formulate and prove a result analogous to Prop. 1.5.13 for nonincreasing sequences.

Exercise 1.5.10.

i.) Let $\{x_k\}$ be a sequence of nonnegative real numbers. For every $n \in \mathbb{N}$, define $s_n = \sum_{k=1}^{n} x_k$. Prove that the sequence $\{s_n\}$ converges iff it is bounded above.

ii.) Prove that $\left\{\frac{a^n}{n!}\right\}$ converges, $\forall a \in \mathbb{R}$. In fact, $\lim_{n\to\infty} \frac{a^n}{n!} = 0$.

Now that we know about subsequences, it is convenient to introduce a generalization of the notion of the limit of a real-valued sequence. We make the following definition.

Definition 1.5.14. Let $\{x_n\} \subset \mathbb{R}$ be a sequence.

i.) We say that $\lim_{n\to\infty} x_n = \infty$ whenever $\forall r > 0$, $\exists N \in \mathbb{N}$ such that $x_n > r$, $\forall n \geq N$. (In such case, we write $x_n \to \infty$.)

ii.) For every $n \in \mathbb{N}$, let $u_n = \sup\{x_k : k \geq n\}$. We define the *limit superior* of x_n by

$$\overline{\lim_{n\to\infty}} x_n = \inf\{u_n | n \in \mathbb{N}\} = \inf_{n\in\mathbb{N}} \left(\sup_{k\geq n} x_k \right).$$

iii.) The dual notion of *limit inferior*, denoted $\underline{\lim}_{n\to\infty} x_n$, is defined analogously with sup and inf interchanged in (ii), viz.,

$$\underline{\lim}_{n\to\infty} x_n = \sup_{n\in\mathbb{N}} \left(\inf_{k\geq n} x_k \right).$$

Some properties of limit superior (inferior) are gathered below. The proofs are standard and can be found in standard analysis texts (see **[234]**).

Proposition 1.5.15. (Properties of Limit Superior and Inferior)
i.) $\overline{\lim_{n\to\infty}} x_n = p \in \mathbb{R}$ *iff* $\forall \varepsilon > 0$,
  a.) *There exist only finitely many n such that* $x_n > p + \varepsilon$, *and*
  b.) *There exist infinitely many n such that* $x_n > p - \varepsilon$;
ii.) $\overline{\lim_{n\to\infty}} x_n = p \in \mathbb{R}$ *iff p is the largest limit of any subsequence of* $\{x_n\}$;
iii.) $\overline{\lim_{n\to\infty}} x_n = \infty$ *iff* $\forall r \in \mathbb{R}$, $\exists$ *infinitely many n such that* $x_n > r$;
iv.) *If* $x_n < y_n$, $\forall n \in \mathbb{N}$, *then*
  a.) $\overline{\lim_{n\to\infty}} x_n \leq \overline{\lim_{n\to\infty}} y_n$,
  b.) $\underline{\lim_{n\to\infty}} x_n \leq \underline{\lim_{n\to\infty}} y_n$;
v.) $\overline{\lim_{n\to\infty}} (-x_n) = -\underline{\lim_{n\to\infty}} x_n$;
vi.) $\underline{\lim_{n\to\infty}} x_n \leq \overline{\lim_{n\to\infty}} x_n$;
vii.) $\lim_{n\to\infty} x_n = p$ *iff* $\underline{\lim_{n\to\infty}} x_n \leq \overline{\lim_{n\to\infty}} x_n = p$;
viii.) $\underline{\lim_{n\to\infty}} x_n + \underline{\lim_{n\to\infty}} y_n \leq \underline{\lim_{n\to\infty}} (x_n + y_n) \leq \overline{\lim_{n\to\infty}} (x_n + y_n) \leq \overline{\lim_{n\to\infty}} x_n + \overline{\lim_{n\to\infty}} y_n$;
ix.) *If* $x_n \geq 0$ *and* $y_n \geq 0$, $\forall n \in \mathbb{N}$, *then* $\overline{\lim_{n\to\infty}} (x_n y_n) \leq \left(\overline{\lim_{n\to\infty}} x_n\right)\left(\overline{\lim_{n\to\infty}} y_n\right)$, *provided the product on the right is not of the form* $0 \cdot \infty$.

1.5.3 Cauchy Sequences

Definition 1.5.16. A sequence $\{x_n\}$ is a *Cauchy sequence* if $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that

$$n, m \geq N \implies |x_n - x_m| < \varepsilon.$$

Intuitively, the terms of a Cauchy sequence squeeze together as the index increases. Given any positive error tolerance ε, there is an index past which any two terms of the sequence, no matter how greatly their indices differ, have values within the tolerance of ε of one another. For brevity, we often write "$\{x_n\}$is Cauchy" instead of "$\{x_n\}$ is a Cauchy sequence."

Exercise 1.5.11. Prove that the following statements are equivalent:
  (1) $\{x_n\}$ is a Cauchy sequence.
  (2) $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that $\left|x_{N+p} - x_{N+q}\right| < \varepsilon, \forall p, q \in \mathbb{N}$.
  (3) $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that $|x_n - x_N| < \varepsilon, \forall n \geq N$.
  (4) $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that $\left|x_{N+p} - x_N\right| < \varepsilon, \forall p \in \mathbb{N}$.
  (5) $\lim_{n\to\infty} (x_{n+p} - x_n) = 0, \forall p \in \mathbb{N}$.

We could have included the statement "$\{x_n\}$ is a convergent sequence" in the above list and asked which others imply it or are implied by it. Indeed, which of the two statements

$$\{x_n\} \text{ is a convergent sequence}$$

or

$$\{x_n\} \text{ is a Cauchy sequence}$$

seems stronger to you? Which implies which, if either? We will revisit this question after the following lemma.

Lemma 1.5.17. (Properties of Cauchy Sequences in $\mathbb{R}$)
i.) *A Cauchy sequence is bounded.*
ii.) *If a Cauchy sequence $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ that converges to L, then $\{x_n\}$ itself converges to L.*

Outline of Proof:
Proof of (i): Let $\{x_n\}$ be a Cauchy sequence. Then, by Def. 1.5.16, $\exists N \in \mathbb{N}$ such that

$$n,m \geq N \Longrightarrow |x_n - x_m| < 1.$$

In particular,

$$n \geq N \Longrightarrow |x_n - x_N| < 1.$$

Starting with the last statement, argue as in Prop. 1.5.6 that $|x_n| \leq M$, $\forall n \in \mathbb{N}$, where

$$M = \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|, |x_N| + 1\}.$$

So, $\{x_n\}$ is bounded.

Proof of (ii): Let $\varepsilon > 0$. $\exists N_1 \in \mathbb{N}$ such that

$$n,m \geq N_1 \Longrightarrow |x_n - x_m| < \frac{\varepsilon}{2} \tag{1.36}$$

and $\exists N_2 \in \mathbb{N}$ such that

$$k \geq N_2 \Longrightarrow \left|x_{n_k} - L\right| < \frac{\varepsilon}{2}. \tag{1.37}$$

(Why?) Now, how do you select N so that (1.36) and (1.37) hold simultaneously? Let $n > N$ and choose any $k \in \mathbb{N}$ such that $k \geq N$. Then, $n \geq N_1$ and $n_k \geq N$. (Why?) As such,

$$n \geq N \Longrightarrow |x_n - L| = \left|x_n - x_{n_k} + x_{n_k} - L\right| \leq \left|x_n - x_{n_k}\right| + \left|x_{n_k} - L\right| < \varepsilon.$$

(Why?) This completes the proof. □

We now shall prove that convergence and Cauchy are equivalent notions in $\mathbb{R}$.

Theorem 1.5.18. (Cauchy Criterion in $\mathbb{R}$)
$\{x_n\}$is convergent $\Longleftrightarrow$ $\{x_n\}$is a Cauchy sequence.

Outline of Proof:
Proof of $\Longrightarrow$): Suppose that $\lim_{n\to\infty} x_n = L$ and let $\varepsilon > 0$. Then, $\exists N \in \mathbb{N}$ such that

$$n \geq N \Longrightarrow |x_n - L| < \frac{\varepsilon}{2}.$$

Thus,

$$n, m \geq N \Longrightarrow |x_n - x_m| \leq |x_n - L| + |x_m - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \text{ (Why?)}$$

Thus, $\{x_n\}$is a Cauchy sequence.
Proof of $\Longleftarrow$): Let $\{x_n\}$be a Cauchy sequence. Then, $\{x_n\}$ is bounded (Why?) and so contains a convergent subsequence $\{x_{n_k}\}$. (Why?) Denote its limit by L. Then, it follows that, in fact, $\{x_n\}$ converges to L, as needed. (Why?) This completes the proof. □

Remark. The proofs of the results:
i.) A bounded sequence has a convergent subsequence,
ii.) A bounded monotone sequence converges,
iii.) Cauchy Criterion in $\mathbb{R}$,
all require the completeness property of $\mathbb{R}$, the first indirectly via Bolzano-Weierstrass, which in turn uses it, the second directly, and the third via use of the first. Actually, all three statements are not only consequences of the completeness property of $\mathbb{R}$, but are equivalent to it. In fact, in more general settings in which order is no longer available (cf. Sections 1.6 and 1.7), completeness of the space is *defined* to be the property that all Cauchy sequences converge in the space.

Exercise 1.5.12. For every $n \in \mathbb{N}$, define $s_n = \sum_{k=1}^{n} \frac{1}{k}$. Prove that $\{s_n\}$ diverges.

1.5.4 A Brief Look at Infinite Series

Sequences defined by forming partial sums using terms of a second sequence (e.g., see Exer. 1.5.12) often arise in applied analysis. You might recognize them by the name *infinite series*. We shall provide the bare essentials of this topic below. A thorough treatment can be found in **[236]**.

Definition 1.5.19. Let $\{a_n\}$ be a sequence in $\mathbb{R}$.
i.) The sequence $\{s_n\}$ defined by $s_n = \sum_{k=1}^{n} a_k$, $n \in \mathbb{N}$ is the *sequence of partial sums of* $\{a_n\}$.
ii.) The pair $(\{a_n\}, \{s_n\})$ is called an *infinite series,* denoted by $\sum_{n=1}^{\infty} a_n$ or $\sum a_n$.
iii.) If $\lim_{n\to\infty} s_n = s$, then we say $\sum a_n$ *converges* and has *sum* s; we write $\sum a_n = s$. Otherwise, we say $\sum a_n$ *diverges*.

Remarks.
1. The sequence of partial sums can begin with an index n strictly larger than 1.

2. Suppose $\sum_{k=1}^{\infty} a_k = s$ and $s_n = \sum_{k=1}^{n} a_k$. Observe that

$$s_n + \underbrace{\sum_{k=n+1}^{\infty} a_k}_{\text{Tail}} = s. \tag{1.38}$$

Because $\lim\limits_{n\to\infty} s_n = s$, it follows that $\lim\limits_{n\to\infty} \sum_{k=n+1}^{\infty} a_k = 0$. (Why?)

Example. (Geometric Series)
Consider the series $\sum_{k=0}^{\infty} cx^k$, where $c, x \in \mathbb{R}$. For every $n \geq 0$, subtracting the expressions for s_n and xs_n yields

$$\begin{aligned} s_n &= c\left[1 + x + x^2 + \ldots + x^n\right] & (1.39)\\ \underline{-xs_n} &= \underline{c\left[x + x^2 + \ldots + x^n + x^{n+1}\right]} & (1.40)\\ (1-x)s_n &= c\left[1 - x^{n+1}\right]. \end{aligned}$$

Hence,

$$s_n = \begin{cases} \frac{c\left[1-x^{n+1}\right]}{1-x}, & x \neq 1, \\ c(n+1), & x = 1. \end{cases}$$

If $|x| < 1$, then $\lim\limits_{n\to\infty} x^{n+1} = 0$, so that $\lim\limits_{n\to\infty} s_n = \frac{c}{1-x}$. Otherwise, $\lim\limits_{n\to\infty} s_n$ does not exist. (Tell why.)

Exercise 1.5.13. Let $p \in \mathbb{N}$. Determine the values of x for which $\sum_{n=p}^{\infty} c(5x+1)^{3n}$ converges, and for such x, determine its sum.

Proposition 1.5.20. *$\sum a_n$ converges iff $\{s_n\}$ is Cauchy iff $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $n \geq N \implies \left|\sum_{k=1}^{p} a_{n+k}\right| < \varepsilon$.*

Outline of Proof: The first equivalence is immediate (Why?) and the second follows from Exer. 1.5.11. (Tell how.) □

Corollary 1.5.21. **(n^{th} -term test)** *If $\sum a_n$ converges, then $\lim\limits_{n\to\infty} a_n = 0$.*

Outline of Proof: Take $p = 1$ in Prop. 1.5.20. □

Exercise 1.5.14. Prove that $\sum_{n=1}^{\infty} \frac{n!}{a^n}$ diverges, $\forall a > 0$.

Proposition 1.5.22. (Comparison Test)
If $a_n, b_n \geq 0$, $\forall n \in \mathbb{N}$, and $\exists c > 0$ and $N \in \mathbb{N}$ such that $a_n \leq cb_n$, $\forall n \geq N$, then
i.) *$\sum b_n$ converges $\Longrightarrow \sum a_n$ converges;*
ii.) *$\sum a_n$ diverges $\Longrightarrow \sum b_n$ diverges.*

Outline of Proof: Use Prop. 1.5.13 (Tell how.) □

Example. Consider the series $\sum_{n=1}^{\infty} \frac{5n}{3^n}$. Because $\lim_{n\to\infty} \frac{5n}{3^{n/2}} = 0$, $\exists N \in \mathbb{N}$ such that

$$n \geq N \Longrightarrow \frac{5n}{3^{n/2}} < 1 \Longrightarrow \frac{5n}{3^n} < \frac{1}{3^{n/2}} = \left(\frac{1}{\sqrt{3}}\right)^n. \tag{1.41}$$

(Why?) But, $\sum_{n=1}^{\infty} \left(\frac{1}{\sqrt{3}}\right)^n$ is a convergent geometric series. Thus, Prop. 1.5.22 implies that $\sum_{n=1}^{\infty} \frac{5n}{3^n}$ converges.

Definition 1.5.23. A series $\sum a_n$ is *absolutely convergent* if $\sum |a_n|$ converges.

It can be shown that rearranging the terms of an absolutely convergent series does not affect convergence (see **[236]**). So, we can regroup terms at will, which is especially useful when groups of terms simplify nicely.

Proposition 1.5.24. (Ratio Test)
Suppose $\sum a_n$ is a series with $a_n \neq 0, \forall n \in \mathbb{N}$. Let

$$r = \varliminf_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \text{ and } R = \varlimsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|,$$

(where R could be ∞). Then,
i.) *$R < 1 \Longrightarrow \sum a_n$ converges absolutely;*
ii.) *$r > 1 \Longrightarrow \sum a_n$ diverges;*
iii.) *If $r \leq 1 \leq R$, then the test is inconclusive.*

Outline of Proof: We argue as in **[234]**.
Proof of (i): Assume $R < 1$ and choose x such that $R < x < 1$. Observe that

$$\begin{aligned} \varlimsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = R < x \Longrightarrow \exists N \in \mathbb{N} \text{ such that } \left|\frac{a_{n+1}}{a_n}\right| \leq x, \forall n \geq N \\ \Longrightarrow |a_{n+1}| \leq |a_n| x, \forall n \geq N. \end{aligned} \tag{1.42}$$

Thus,

$$\begin{aligned} |a_{N+1}| &\leq |a_N| x \\ |a_{N+2}| &\leq |a_{N+1}| x \leq |a_N| x^2 \\ |a_{N+3}| &\leq |a_{N+2}| x \leq |a_{N+1}| x^2 \leq |a_N| x^3 \\ &\vdots \end{aligned}$$

(Why?) What can be said about the series

$$|a_N| \left(x + x^2 + x^3 + \ldots\right)?$$

Use Prop. 1.5.22 to conclude that $\sum |a_n|$ converges.

Proof of (ii): Next, assume $1 < r$. Observe that

$$\begin{aligned}\varliminf_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = r > 1 &\Longrightarrow \exists N \in \mathbb{N} \text{ such that } \left|\frac{a_{n+1}}{a_n}\right| \geq 1, \forall n \geq N \\ &\Longrightarrow |a_{n+1}| \geq |a_n| \geq |a_N| > 0, \forall n \geq N. \end{aligned} \tag{1.43}$$

(Why?) Thus, $a_n \nrightarrow 0$. (So what?)

Proof of (iii): For both $\sum \frac{1}{n}$ and $\sum \frac{1}{n^2}$, $r = R = 1$, but $\sum \frac{1}{n}$ diverges and $\sum \frac{1}{n^2}$ converges. □

Exercise 1.5.15. Determine if $\sum_{n=1}^{\infty} \frac{n^n}{n!}$ converges.

Finally, we will need to occasionally multiply two series in the following sense.

Definition 1.5.25. Given two series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$, define

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}, \forall n \geq 0.$$

The series $\sum_{n=0}^{\infty} c_n$ is called the *Cauchy product* of $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$.

To see why this is a natural definition, consider the partial sum $\sum_{n=0}^{p} c_n$ and form a grid by writing the terms $a_0, \ldots, a_p$ as a column and $b_0, \ldots, b_p$ as a row. Multiply the terms from each row and column pairwise and observe that the sums along the diagonals (formed left to right) coincide with $c_0, \ldots, c_p$. (Check this!)

The following proposition describes a situation when such a product converges. The proof of this and other related results can be found in **[17]**.

Proposition 1.5.26. *If $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ both converge absolutely, then the Cauchy product $\sum_{n=0}^{\infty} c_n$ converges absolutely and $\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right)$.*

1.6 The Spaces $\left(\mathbb{R}^N, \|\cdot\|_{\mathbb{R}^N}\right)$ and $\left(\mathbb{M}^N(\mathbb{R}), \|\cdot\|_{\mathbb{M}^N(\mathbb{R})}\right)$

We now introduce two spaces of objects with which you likely have some familiarity, namely vectors and square matrices, as a first step in formulating more abstract spaces. The key observation is that the characteristic properties of $\mathbb{R}$ carry over to these spaces, and their verification requires minimal effort. As you work through this section, use your intuition about how vectors in two and three dimensions behave to help you understand the more abstract setting.

1.6.1 The Space $\left(\mathbb{R}^N, \|\cdot\|_{\mathbb{R}^N}\right)$

Definition 1.6.1. For every $N \in \mathbb{N}$, $\mathbb{R}^N = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{N\,\text{times}}$ is the set of all ordered N-tuples of real numbers. This set is often loosely referred to as *N-space*.

A typical element of $\mathbb{R}^N$ (called a *vector*) is denoted by a boldface letter, say $\mathbf{x}$, representing the ordered N-tuple $\langle x_1, x_2, \ldots, x_N \rangle$. (Here, x_k is the k^{th} *component* of $\mathbf{x}$.) The *zero element* in $\mathbb{R}^N$ is the vector $\mathbf{0} = \underbrace{\langle 0, 0, \ldots, 0 \rangle}_{N\,\text{times}}$.

The algebraic operations defined in $\mathbb{R}$ can be applied componentwise to define the corresponding operations in $\mathbb{R}^N$. Indeed, we have

Definition 1.6.2. (Algebraic Operations in $\mathbb{R}^N$)
Let $\mathbf{x} = \langle x_1, x_2, \ldots, x_N \rangle$ and $\mathbf{y} = \langle y_1, y_2, \ldots, y_N \rangle$ be elements of $\mathbb{R}^N$ and $c \in \mathbb{R}$,
i.) $\mathbf{x} = \mathbf{y}$ if and only if $x_k = y_k$, $\forall k \in \{1, \ldots, N\}$,
ii.) $\mathbf{x} + \mathbf{y} = \langle x_1 + y_1, x_2 + y_2, \ldots, x_N + y_N \rangle$,
iii.) $c\mathbf{x} = \langle cx_1, cx_2, \ldots, cx_N \rangle$.

The usual properties of commutativity, associativity, and distributivity of scalar multiplication over addition carry over to this setting by applying the corresponding property in $\mathbb{R}$ componentwise. For instance, because $x_i + y_i = y_i + x_i$, $\forall i \in \{1, \ldots, n\}$, it follows that

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= \langle x_1 + y_1, x_2 + y_2, \ldots, x_N + y_N \rangle \\ &= \langle y_1 + x_1, y_2 + x_2, \ldots, y_N + x_N \rangle \\ &= \mathbf{y} + \mathbf{x}. \end{aligned} \tag{1.44}$$

Exercise 1.6.1. Establish associativity of addition and distributivity of scalar multiplication over addition in $\mathbb{R}^N$.

Geometric and Topological Structure

From the viewpoint of its geometric structure, what is a natural candidate for a distance-measuring artifice for $\mathbb{R}^N$? There is more than one answer to this question, arguably the most natural of which is the Euclidean distance formula, defined below.

Definition 1.6.3. Let $\mathbf{x} \in \mathbb{R}^N$. The *(Euclidean) norm* of $\mathbf{x}$, denoted $\|\mathbf{x}\|_{\mathbb{R}^N}$, is defined by

$$\|\mathbf{x}\|_{\mathbb{R}^N} = \sqrt{\sum_{k=1}^{N} x_k^2}. \tag{1.45}$$

We say that the *distance between $\mathbf{x}$ and $\mathbf{y}$ in* $\mathbb{R}^N$ is given by $\|\mathbf{x} - \mathbf{y}\|_{\mathbb{R}^N}$.

Remarks.
1. When referring to the norm generically or as a function, we write $\|\cdot\|_{\mathbb{R}^N}$.

2. There are other "equivalent" ways to define a norm on $\mathbb{R}^N$ that are more convenient to use in some situations. Indeed, a useful alternative norm is given by

$$\|\mathbf{x}\|_{\mathbb{R}^N} = \max_{1\leq i\leq N} |x_i| \,. \tag{1.46}$$

By *equivalent*, we do not mean that the numbers produced by (1.45) and (1.46) are the same for a given $\mathbf{x}\in\mathbb{R}^N$. In fact, this is false in a big way! Rather, two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are *equivalent* if there exist constants $0<\alpha<\beta$ such that

$$\alpha\,\|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2 \leq \beta\,\|\mathbf{x}\|_1\,, \forall \mathbf{x}\in\mathbb{R}. \tag{1.47}$$

Suffice it to say that you can choose whichever norm is most convenient to work with within a given series of computations, as long as you don't decide to use a different one halfway through! **By default, we use (1.45) unless otherwise specified.**

Exercise 1.6.2. Let $\varepsilon>0$. Provide a geometric description of these sets:
i.) $A=\left\{\mathbf{x}\in\mathbb{R}^2 \mid \|\mathbf{x}\|_{\mathbb{R}^2}<\varepsilon\right\}$,
ii.) $B=\left\{\mathbf{y}\in\mathbb{R}^3 \mid \|\mathbf{y}-\langle 1,0,0\rangle\|_{\mathbb{R}^3}\geq\varepsilon\right\}$,
iii.) $C=\left\{\mathbf{y}\in\mathbb{R}^3 \mid \|\mathbf{y}-\mathbf{x}_0\|_{\mathbb{R}^3}=0\right\}$, where $\mathbf{x}_0\in\mathbb{R}^3$ is prescribed.

The $\mathbb{R}^N$-norm satisfies similar properties as $|\cdot|$ (cf. Prop. 1.4.4), summarized below.

Proposition 1.6.4. *Let* $\mathbf{x},\mathbf{y}\in\mathbb{R}^N$ *and* $c\in\mathbb{R}$. *Then,*
i.) $\|\mathbf{x}\|_{\mathbb{R}^N}\geq 0$,
ii.) $\|c\mathbf{x}\|_{\mathbb{R}^N}=|c|\,\|\mathbf{x}\|_{\mathbb{R}^N}$,
iii.) $\|\mathbf{x}+\mathbf{y}\|_{\mathbb{R}^N}\leq\|\mathbf{x}\|_{\mathbb{R}^N}+\|\mathbf{y}\|_{\mathbb{R}^N}$,
iv.) $\mathbf{x}=\mathbf{0}$ *iff* $\|\mathbf{x}\|_{\mathbb{R}^N}=0$.

Exercise 1.6.3. Prove Prop. 1.6.4 using Def. 1.6.3. Then, redo it using (1.46).

Exercise 1.6.4. Let $M,\, p\in\mathbb{N}$. Prove the following string of inequalities:

$$\left\|\sum_{i=1}^{M}\mathbf{x}_i\right\|_{\mathbb{R}^N}^{p} \leq \left(\sum_{i=1}^{M}\|\mathbf{x}_i\|_{\mathbb{R}^N}\right)^{p} \leq M^{p-1}\sum_{i=1}^{M}\|\mathbf{x}_i\|_{\mathbb{R}^N}^{p} \tag{1.48}$$

The space $\left(\mathbb{R}^N, \|\cdot\|_{\mathbb{R}^N}\right)$ has an even richer geometric structure since it can be equipped with a so-called *inner product* that enables us to define orthonormality (or perpendicularity) and, by extension, the notion of angle in the space. Precisely, we have

Definition 1.6.5. Let $\mathbf{x},\mathbf{y}\in\mathbb{R}^N$. The *inner product of* $\boldsymbol{x}$ *and* $\boldsymbol{y}$, denoted $\langle\mathbf{x},\mathbf{y}\rangle_{\mathbb{R}^N}$, is defined by

$$\langle\mathbf{x},\mathbf{y}\rangle_{\mathbb{R}^N}=\sum_{i=1}^{N}x_iy_i. \tag{1.49}$$

Note that taking the inner product of any two elements of $\mathbb{R}^N$ produces a real number. Also, $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}^N}$ is often written more compactly as $\mathbf{x}\mathbf{y}^{\mathrm{T}}$, where $\mathbf{y}^{\mathrm{T}}$ is the transpose of $\mathbf{y}$ (that is, $\mathbf{y}$ written as a column vector rather than as a row vector). Some of the properties of this inner product are as follows.

Proposition 1.6.6. (Properties of the Inner Product on $\mathbb{R}^N$)
Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^N$ *and* $c \in \mathbb{R}$. *Then,*
i.) $\langle c\mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}^N} = \langle \mathbf{x}, c\mathbf{y} \rangle_{\mathbb{R}^N} = c \langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}^N}$;
ii.) $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle_{\mathbb{R}^N} = \langle \mathbf{x}, \mathbf{z} \rangle_{\mathbb{R}^N} + \langle \mathbf{y}, \mathbf{z} \rangle_{\mathbb{R}^N}$;
iii.) $\langle \mathbf{x}, \mathbf{x} \rangle_{\mathbb{R}^N} \geq 0$;
iv.) $\langle \mathbf{x}, \mathbf{x} \rangle_{\mathbb{R}^N} = 0$ *iff* $\mathbf{x} = \mathbf{0}$;
v.) $\langle \mathbf{x}, \mathbf{x} \rangle_{\mathbb{R}^N} = \|\mathbf{x}\|^2_{\mathbb{R}^N}$;
vi.) $\langle \mathbf{x}, \mathbf{z} \rangle_{\mathbb{R}^N} = \langle \mathbf{y}, \mathbf{z} \rangle_{\mathbb{R}^N}$, $\forall \mathbf{z} \in \mathbb{R}^N \Longrightarrow \mathbf{x} = \mathbf{y}$.

Verifying these properties is straightforward and will be argued in a more general setting in Section 1.7. (Try proving them here!) Property (v) is of particular importance because it asserts that an inner product generates a norm.

Exercise 1.6.5. Prove Prop. 1.6.6.

The following Cauchy-Schwarz inequality is very important.

Proposition 1.6.7. (Cauchy-Schwarz Inequality)
Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$. *Then,*

$$|\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}^N}| \leq \|\mathbf{x}\|_{\mathbb{R}^N} \|\mathbf{y}\|_{\mathbb{R}^N} \tag{1.50}$$

Outline of Proof: For any $\mathbf{y} \in \mathbb{R}^N \setminus \{\mathbf{0}\}$,

$$0 \leq \left\langle \mathbf{x} - \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}^N}}{\|\mathbf{y}\|^2_{\mathbb{R}^N}} \right) \mathbf{y}, \mathbf{x} - \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}^N}}{\|\mathbf{y}\|^2_{\mathbb{R}^N}} \right) \mathbf{y} \right\rangle_{\mathbb{R}^N}.$$

(So what?) Why does (1.50) hold for $\mathbf{y} = \mathbf{0}$? □

The inner product can be used to formulate a so-called *orthonormal basis* for $\mathbb{R}^N$. Precisely, let

$$\mathbf{e}_1 = \langle 1, 0, \ldots, 0 \rangle, \mathbf{e}_2 = \langle 0, 1, 0, \ldots, 0 \rangle, \ldots, \mathbf{e}_n = \langle 0, \ldots, 0, 1 \rangle,$$

and observe that

$$\|\mathbf{e}_i\|_{\mathbb{R}^N} = 1, \forall i \in \{1, \ldots, N\}, \tag{1.51}$$

$$\left\langle \mathbf{e}_i, \mathbf{e}_j \right\rangle_{\mathbb{R}^N} = 0, \text{ whenever } i \neq j. \tag{1.52}$$

This is useful because it yields the following unique representation for the members of $\mathbb{R}^N$ involving the inner product.