Lecture Notes in Real Analysis

Eric T. Sawyer

McMaster University, Hamilton, Ontario
E-mail address: sawyer@mcmaster.ca

Abstract. Beginning with the ordered field of real numbers, these lecture notes examine the theory of real functions with applications to differential equations and fractals. The main thread begins with the least upper bound property of the real numbers, and follows through to compactness and completeness in Euclidean spaces. Standard results on continuity, differentiation and integration are established, culminating in two applications of the Contraction Lemma: fractals are characterized using the completeness of the metric space of compact subsets of Euclidean space; existence and uniqueness of solutions to first order nonlinear initial value problems are proved using completeness of the space of real continuous functions on a closed bounded interval.

Contents

Preface

These notes grew out of lectures given three times a week in a third year undergraduate course in real analysis at McMaster University September to December 2009. The topics include the real and complex number systems and their function theory; continuity, differentiability, and compactness. Applications include existence of solutions to differential equations, and constuction of fractals such as the Cantor set, the von Koch snowflake and Peano's space-filling curve. Sources include books by Rudin [**3**] and [**4**], books by Stein and Shakarchi [**5**] and [**6**], and the history book by Boyer [**2**].

Part 1

Differentiation

We begin Part 1 with a chapter discussing the field of real numbers $\mathbb{R}$, in particular its status as the *unique* ordered field with the least upper bound property. We show that the field of real numbers $\mathbb{R}$ can be constructed either from *Dedekind cuts* of rational numbers $\mathbb{Q}$, or from Weierstrass' *Cauchy sequences* of rational numbers. Finally, we comment briefly on the arithmetic properties of $\mathbb{R}$ that can be derived from its definition, and also point out a false start in the construction of $\mathbb{R}$.

Then in the short Chapter 2 we introduce Cantor's cardinal numbers and show that the rational numbers are countable and that the real numbers are uncountable.

Chapter 3 follows Rudin [**3**] in part and introduces the concept of a metric space with a 'distance function' that is sufficient for developing a rich theory of limits, yet general enough to include the real and complex numbers, Euclidean spaces and the various function spaces we use later. We also construct our first fractal set, the famous Cantor middle thirds set, which provides an example of a perfect set that is large in cardinality (uncountable) yet small in 'length' (measure zero). We end by following Stein and Sharkarchi [**6**] to establish a one-to-one correspondence between finite collections of contractive similarities and fractal sets, thus illustrating Mandelbrot's observation that much of the apparent chaotic form in nature has an extremely simple underlying structure.

Chapter 4 develops the standard theory of sequences and series in a metric space, including convergence tests, Cauchy sequences and the completeness of Euclidean spaces. We also introduce the useful contraction lemma as a unifying approach to fractals and later, to solutions to differential equations.

Chapter 5 introduces the concepts of continuity and differentiability including uniform continuity and four mean value theorems of increasing generality.

CHAPTER 1

The fields of analysis

If one is not careful in defining the concepts used in analysis, confusion can result. In particular, we need a clear definition of

(1) **function**,
(2) the set of **real numbers**, and
(3) **convergence** of series of real numbers and functions.

In the 18^{th} century each of these concepts suffered shortcomings. Early formulations of the notion of function involved the idea of a specific formula. Later in 1837, Lejeune Dirichlet suggested a broader definition of function, still falling short of the modern notion:

- If a variable y is so related to a variable x that whenever a numerical value is assigned to x, there is a *rule* according to which a unique value of y is determined, then y is said to be a function of the independent variable x.

Real numbers were thought of as points on a line, but the identification of their crucial properties, such as the absence of gaps as reflected in the least upper bound property, had to await Dedekind's construction of the real numbers from the rational numbers.

In 1725 Varignon, one of the first French scholars to appreciate the calculus, warned that infinite series were not to be used without investigation of the remainder term. It was not until 1872 however before Heine, influenced by Weierstrass' lectures, defined the limit of the function f at x_0 in virtually modern terms as follows:

- If, given any ε, there is an η_0 such that for $0 < \eta < \eta_0$ the difference $f(x_0 \pm \eta) - L$ is less in absolute value than ε, then L is the limit of $f(x)$ for $x = x_0$.

Historically, the following example was pivotal in the development of the rigorous analysis that addressed the above shortcomings, and also in the foundations of set theory. We are referring here to a simple mathematical model of the motion of a string vibrating in the plane.

1. A model of a vibrating string

Consider a vibrating string stretched along that portion of the x-axis in the plane that joins the points $(0,0)$ and $(1,0)$, and suppose the string is wiggling up and down (not very violently) in the y-direction. Suppose that at time t and just above (or below) the point $(x,0)$ on the x-axis, the y-coordinate of the string is given by $y(x,t)$. This defines a 'function' mapping the infinite strip $[0,1] \times \mathbb{R}$ into the real numbers $\mathbb{R}$, i.e. $y(x,t)$ is defined for

$$0 \leq x \leq 1 \text{ and } t \in \mathbb{R},$$

and we are to think of the real number $y(x,t)$ as measuring the displacement from the x-axis of the vibrating string at position x and time t. We assume the endpoints of the string are attached to the points $(0,0)$ and $(1,0)$ for all time and so we have the boundary conditions

$$y(0,t) = 0 \text{ and } y(1,t) = 0 \text{ for all } t \in \mathbb{R}. \tag{1.1}$$

Moreover, we can suppose that at time $t = 0$ the shape of the string is specified by the graph of a given function f that maps $[0,1]$ to $\mathbb{R}$;

$$y(x,0) = f(x) \text{ for } 0 \leq x \leq 1. \tag{1.2}$$

Finally, we can suppose that at time $t = 0$ the vertical velocity of the string is specified by a given function g that maps $[0,1]$ to $\mathbb{R}$;

$$\frac{\partial}{\partial t} y(x,0) = g(x) \text{ for } 0 \leq x \leq 1. \tag{1.3}$$

Now provided the displacements are not too violent, it can be shown (and we are not interested here in exactly how this is done) that the function $y(x,t)$ satisfies a partial differential equation of the form

$$\frac{\partial^2}{\partial t^2} y = c^2 \frac{\partial^2}{\partial x^2} y, \qquad 0 < x < 1 \text{ and } t \in \mathbb{R},$$

where c is a positive constant determined by the physical properties of the string, and is interpreted as the speed of propagation. This is the so-called wave equation, and together with the boundary conditions (1.1) and the initial conditions (1.2) and (1.3), it constitutes the *initial boundary value problem* for the vibrating string:

$$\begin{aligned} \left(\frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial x^2}\right) y(x,t) &\equiv 0, \qquad 0 < x < 1 \text{ and } t \in \mathbb{R}, \\ y(0,t) &= y(1,t) = 0, \qquad t \in \mathbb{R}, \\ \begin{bmatrix} y(x,0) \\ \frac{\partial}{\partial t} y(x,0) \end{bmatrix} &= \begin{bmatrix} f(x) \\ g(x) \end{bmatrix}, \qquad 0 \leq x \leq 1. \end{aligned} \tag{1.4}$$

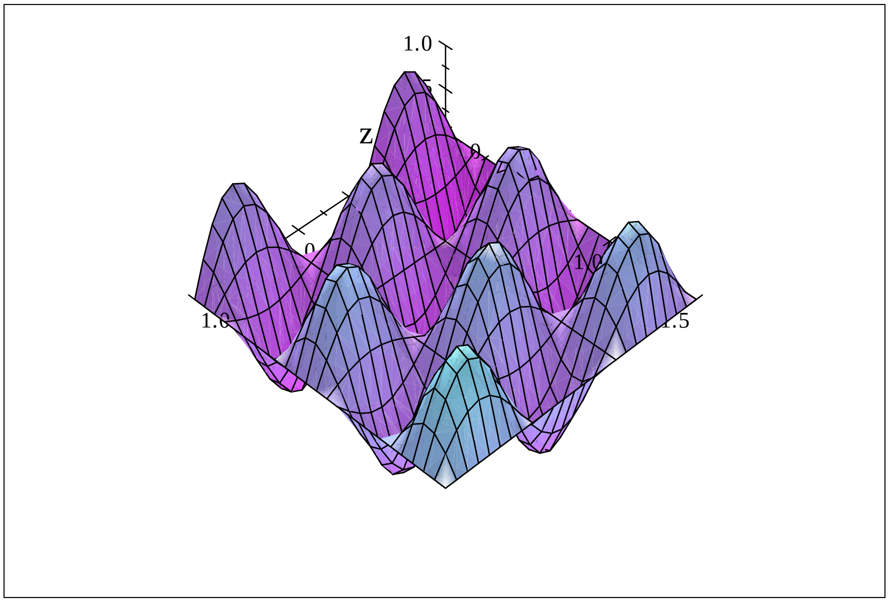
On the one hand, Daniel Bernoulli noted around the middle of the 18^{th} century that for each positive integer $n \in \mathbb{N}$ the function

$$y_n(x,t) = (\sin n\pi x)(\cos nc\pi t),$$

is a solution to (1.4) with initial conditions

$$\begin{aligned} f(x) &= \sin n\pi x, \qquad 0 \leq x \leq 1, \\ g(x) &= 0, \qquad 0 \leq x \leq 1. \end{aligned}$$

EXAMPLE 1. $y(x,t) = (\sin 3\pi x)(\cos 3\pi t)$



Since the equations involved are linear we then have that

$$y(x,t)=\sum_{n=1}^{N}a_n(\sin n\pi x)(\cos nc\pi t)$$

is a solution to (1.4) with initial conditions

$$\begin{aligned} f(x) &= \sum_{n=1}^{N}a_n\sin n\pi x, \qquad 0\le x\le 1,\\ g(x) &= 0, \qquad 0\le x\le 1. \end{aligned}$$

Presuming that we can take infinite sums, we finally obtain that the solution $y(x,t)$ to the initial boundary value problem (1.4) with initial conditions

$$\begin{aligned} f(x) &= \sum_{n=1}^{\infty}a_n\sin n\pi x, \qquad 0\le x\le 1,\\ g(x) &= 0, \qquad 0\le x\le 1, \end{aligned}$$

is given by the infinite series of functions

$$y(x,t)=\sum_{n=1}^{\infty}a_n(\sin n\pi x)(\cos nc\pi t). \tag{1.5}$$

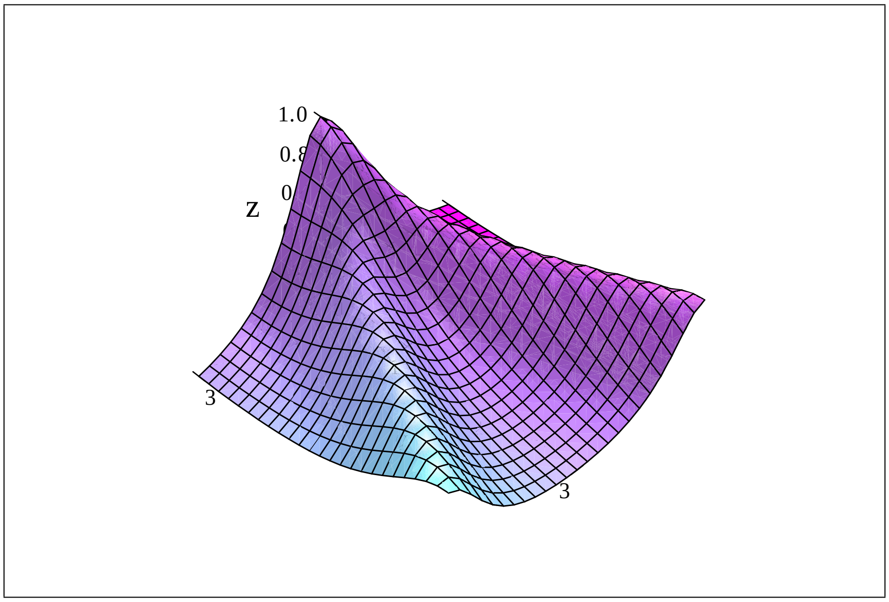
Remark 1. *The Bernoulli decomposition is motivated for example by plucking a guitar string. The fundamental note heard is that corresponding to $n=1$, the standing sine wave having one node that oscillates with frequency $\frac{c}{2}$ and amplitude a_1. Corresponding to higher values of n are the harmonics having n nodes with frequency $\frac{nc}{2}$ and amplitude a_n. See Example 1 above where the standing wave having 3 nodes has graph $\sin 3\pi x$ with frequency $\frac{3}{2}$ and amplitude 1.*

On the other hand, a much simpler solution to (1.4) with initial condition $g(x)=0$ for $0\le x\le 1$ was given by Jean Le Rond d'Alembert in 1747, namely the travelling wave solution ,

$$y(x,t) = \frac{f(x+ct) + f(x-ct)}{2}, \qquad 0 \le x \le 1 \text{ and } t \in \mathbb{R}, \tag{1.6}$$

where we define f outside the interval $[0,1]$ by requiring that it be odd on the interval $[-1,1]$ and periodic with period 2 on the real line.

EXAMPLE 2. $y(x,t) = \frac{1}{2}\frac{1}{1+(x+t)^2} + \frac{1}{2}\frac{1}{1+(x-t)^2}, \qquad -\infty < x < \infty, t \ge 0$



EXERCISE 1. *Verify that the function $y(x,t)$ in (1.6) satisfies (1.4) with $g \equiv 0$.*

REMARK 2. *The travelling wave solution is motivated for example by snapping a skipping rope that is lying in a line on the ground. A 'hump' is produced that travels like a wave along the rope with speed c. See Example 2 above where two 'humps' move off in opposite directions with speed* 1.

Based on physical experience, such as plucking a guitar string and snapping a skipping rope, we expect that

(1) every solution to the initial boundary value problem (1.4) has the Fourier harmonic form (1.5), and
(2) every solution to the initial boundary value problem (1.4) has the d'Alembert travelling wave form (1.6), and
(3) the solution to the initial boundary value problem (1.4) is *uniquely* determined by the boundary conditions (1.1) and the initial conditions (1.2) and (1.3).

From these expectations it follows that for *any* function $f(x)$ we have

$$f(x) = y(x,0) = \sum_{n=1}^{\infty} a_n \sin n\pi x, \qquad 0 \le x \le 1, \tag{1.7}$$

for a suitable choice of constants a_n, $n \ge 1$. The precise meaning to be attached to such a formula (1.7) involves many difficulties! In particular,

- when does the series on the right converge?
- and for what values of x?
- or more generally in what sense?
- and when does the sum equal $f(x)$ in some sense?

We will introduce concepts and develop tools to answer such questions. In particular we note that it was Joseph Fourier in 1824 who first proved that (1.7) holds under certain conditions, and this is the reason that the name of Fourier, and not Bernoulli, is associated with such a decomposition of a function $f(x)$ into a series of trigonometric functions $\sin n\pi x$.

One question that springs to mind immediately is whether or not the ordered field of rational numbers

$$\mathbb{Q} = \left\{ \frac{m}{n} : m \in \mathbb{Z} \text{ and } n \in \mathbb{N} \right\}$$

can suffice as the domain for x in answering these questions. As it happens, the rational numbers suffer a fatal deficiency that we show can morph into different forms in the next section, rendering the rationals unsuitable for this purpose. It is convenient at this point to introduce the concept of an *order* $<$ on a set S.

DEFINITION 1. *An* order $<$ *on a set S is a relation (among ordered pairs (x, y) of elements $x, y \in S$) satisfying the following three properties:*

(1) ***(nonreflexive)*** *If $x \in S$, then it is* not *true that $x < x$.*
(2) ***(antisymmetric)*** *If $x, y \in S$ and $x \neq y$, then* one *and* only one *of the following two possibilities holds:*

$$x < y, \qquad y < x.$$

(3) ***(transitive)*** *If $x, y, z \in S$, and $x < y$ and $y < z$, then $x < z$.*

For example, the usual order on either $\mathbb{Z}$ or $\mathbb{Q}$ satisfies Definition 1.

2. Deficiencies of the rational numbers

The rational numbers $\mathbb{Q}$ form an ordered field, but there are difficulties associtated with

(1) **nonsolvability** of algebraic equations,
(2) **gaps** in the order,
(3) and **nonexistence** of solutions to simple differential equations.

Because of these problems with the rational numbers, we will be led to construct the set of real numbers $\mathbb{R}$ which form an ordered field with the *least upper bound property*. This last property reflects the absence of gaps in the order of the real numbers and accounts for the privileged position of $\mathbb{R}$ in analysis.

2.1. Nonsolvability of algebraic equations. The polynomial equation

$$x^2 - 2 = 0$$

has no solution $x \in \mathbb{Q}$. Indeed, if it did then we would have $\left(\frac{m}{n}\right)^2 = 2$ where m and n are integers with no factors in common. Then

- $m^2 = 2n^2$ is even,
- hence so is m, say $m = 2k$ for an integer k,
- hence $n^2 = 2k^2$ is even,
- and hence n is even.

This contradicts our assumption that m and n have no factors in common, and completes the proof that $\sqrt{2}$ is not rational.

Alternatively, one can avoid divisibility and argue with inequalities to derive a contradiction as Fermat did:

- $\sqrt{2} = \frac{m}{n}$ where $0 < n < m < 2n$,
- $1 = 2 - 1 = \left(\sqrt{2} - 1\right)\left(\sqrt{2} + 1\right) = \left(\frac{m}{n} - 1\right)\left(\sqrt{2} + 1\right)$,
- $\sqrt{2} = \frac{1}{\frac{m}{n} - 1} - 1 = \frac{2n-m}{m-n} = \frac{m_1}{n_1}$ where $n_1 = m - n < n$.

Thus we have shown that if $\sqrt{2}$ can be represented as a quotient of positive integers $\frac{m}{n}$, then it can also be represented as a quotient of positive integers $\frac{m_1}{n_1}$ with n_1 *strictly* smaller than n. This can be repeated as often as we wish, leading to the contradiction that there are infinitely many integers between 1 and n. This technique is known as *Fermat's method of infinite descent*.

REMARK 3. *The equation $x^2 + 2 = 0$ has no solution in $\mathbb{Q}$ either, in fact it has no solution in the real numbers $\mathbb{R}$. This prompts introduction of the set of complex numbers $\mathbb{C}$, which turns out to be an algebraically closed field containing the reals, i.e. every polynomial with real (even complex) coefficients has a root in $\mathbb{C}$. On the other hand, $\mathbb{C}$ is* not *an ordered field, which explains why so much of analysis begins with the real field $\mathbb{R}$.*

2.2. Gaps in the order. The rational numbers can be decomposed into two disjoint sets A and B with the properties that A has no largest element and B has no smallest element, thus leaving a *gap* in the order. By this we mean that we could insert a new element labelled X, @ or even $\sqrt{2}$ into $\mathbb{Q}$ and extend the order on $\mathbb{Q}$ to the larger set $\mathbb{Q} \cup \{X\}$ by declaring $p < X < q$ for all $p \in A$ and $q \in B$. Because this extended order on $\mathbb{Q} \cup \{X\}$ satisfies Definition 1, we say that the sets A and B create a *gap* in the order of $\mathbb{Q}$.

For example we can set

$$\begin{aligned} A &= \left\{p \in \mathbb{Q} : \text{either } p \leq 0 \text{ or } p^2 < 2\right\}, \\ B &= \left\{q \in \mathbb{Q} : q > 0 \text{ and } q^2 > 2\right\}. \end{aligned} \tag{2.1}$$

To see that A has no largest element, pick $p \in A$. We may assume that $p > 0$, and since every p in A is less than 2 we have $0 < p < 2$. Set $\delta = \frac{2-p^2}{8}$ so that $0 < \delta < \frac{1}{4}$. Then

$$\begin{aligned} (p+\delta)^2 &= p^2 + 2p\delta + \delta^2 \\ &< p^2 + 4\delta + \frac{1}{4}\delta \\ &= p^2 + \frac{1}{2}\left(2 - p^2\right) + \frac{1}{32}\left(2 - p^2\right) \\ &< p^2 + \left(2 - p^2\right) = 2. \end{aligned}$$

Thus $p + \delta > p$ and $p \in A$. The proof that B has no smallest element is similar.

2.3. Nonexistence of solutions to differential equations. The differential equation

$$y' + xy^3 = 0$$

has no solution on any open interval of rational numbers. Indeed, we can solve the equation in the real line by separating variables;

$$\begin{aligned} -\frac{1}{2}d\left(\frac{1}{y^2}\right) &= \frac{dy}{y^3} = -xdx = -\frac{1}{2}d\left(x^2\right), \\ \frac{1}{y^2} &= x^2 + C, \\ y &= \frac{1}{\sqrt{x^2+C}}. \end{aligned}$$

No matter what choice of integration constant C is made, and what choice of interval (a, b) with rational numbers $a < b$, there are lots of rational numbers $x \in (a, b)$ for which $y = \frac{1}{\sqrt{x^2+C}}$ is *not* rational.

3. The real field

In regards to the problem of describing what is meant by the 'continuity of a line segment', J. W. R. Dedekind published his famous construction of the real numbers using Dedekind cuts in 1872. Some years earlier he had described his seminal idea in the following way "By this commonplace remark the secret of continuity is to be revealed", the idea in question being

- In any division of the points of the segment into two parts such that each point belongs to one and only one class, and such that every point of the one class is to the left of every point in the other, there is one and only one point that brings about the division.

We present here a modification of this idea due to Bertand Russell (born 1872, the year of Dedekind's publication). Heuristically, following Russell, a Dedekind cut $\alpha \subset \mathbb{Q}$ is a "left infinite interval open on the right" of rational numbers that is associated with the "real number" on the number line that marks its right hand endpoint. More precisely, a cut α is a subset of $\mathbb{Q}$ satisfying (here p and q denote rational numbers)

$$\begin{aligned} \alpha &\neq \emptyset \text{ and } \alpha \neq \mathbb{Q}, \\ p &\in \alpha \text{ and } q < p \text{ implies } q \in \alpha, \\ p &\in \alpha \text{ implies there is } q \in \alpha \text{ with } p < q. \end{aligned} \tag{3.1}$$

One can define an ordered field structure on the set of cuts, which we identify as the field $\mathbb{R}$ of real numbers, and prove that this ordered field has the famous *Least Upper Bound Property* defined below. It is this property that evolves into the critical Heine-Borel property of Euclidean space, namely that every closed and bounded subset is *compact*, and this property in turn ultimately permits the familiar existence theorems for ordinary and partial differential equations. We remark that a copy of the rational number field $\mathbb{Q}$ can be identified inside the real field $\mathbb{R}$ of Dedekind cuts by associating to each $r \in \mathbb{Q}$ the cut

$$\alpha = (-\infty, r) \equiv \{p \in \mathbb{Q} : p < r\}.$$

Alternatively, one can define an ordered field structure on the set of equivalence classes of *Cauchy sequences* in $\mathbb{Q}$, and this produces an ordered field isomorphic to $\mathbb{R}$. We will construct the real numbers using Dedekind cuts at the end of this chapter, and leave the construction with Cauchy sequences to a later chapter. But first we study some of the consequences of an ordered field with the least upper bound property. For this we introduce precise definitions of these concepts.

DEFINITION 2. *A* field $\mathbb{F}$ *is a set with two binary operations, called* addition *and* multiplication*, that satisfy the following three sets of axioms. We often write* $\mathbb{F}$ *for the underlying set,* $x+y$ *for the operation of addition applied to* $x, y \in \mathbb{F}$*, and juxtaposition* xy *for the operation of multiplication applied to* $x, y \in \mathbb{F}$*.*

(1) ***Addition Axioms***
  (a) *(closure)* $x+y \in \mathbb{F}$ *for all* $x, y \in \mathbb{F}$*,*
  (b) *(commutativity)* $x+y = y+x$ *for all* $x, y \in \mathbb{F}$*,*
  (c) *(associativity)* $(x+y)+z = x+(y+z)$ *for all* $x, y, z \in \mathbb{F}$*,*
  (d) *(additive identity) There is an element* $0 \in \mathbb{F}$ *such that*
$$0 + x = x \text{ for all } x \in \mathbb{F},$$
  (e) *(inverses) For each* $x \in \mathbb{F}$ *there is an element* $-x \in \mathbb{F}$ *such that*
$$x + (-x) = 0.$$
(2) ***Multiplication Axioms***
  (a) *(closure)* $xy \in \mathbb{F}$ *for all* $x, y \in \mathbb{F}$*,*
  (b) *(commutativity)* $xy = yx$ *for all* $x, y \in \mathbb{F}$*,*
  (c) *(associativity)* $(xy)\,z = x\,(yz)$ *for all* $x, y, z \in \mathbb{F}$*,*
  (d) *(multiplicative identity) There is an element* $1 \in \mathbb{F}$ *such that*
$$1x = x \text{ for all } x \in \mathbb{F},$$
  (e) *(inverses) For each* $x \in \mathbb{F} \setminus \{0\}$ *there is an element* $\frac{1}{x} \in \mathbb{F}$ *such that*
$$x\left(\frac{1}{x}\right) = 1.$$
(3) ***Distributive Law***
$$x\,(y+z) = xy + xz$$
  for all $x, y, z \in \mathbb{F}$*.*

EXAMPLE 3. *The set of rational numbers* $\mathbb{Q}$ *is a field with the usual operations of addition and multiplication. Another example is given by the finite set of integers*
$$\mathbb{F}_p = \{0, 1, 2, ..., p-1\},$$
with addition and multiplication defined modulo p*. This turns out to be a field if and only if* p *is a prime number. Details are left to the reader.*

All of the familiar algebraic identities that hold for the rational numbers, hold also in any field. We state the most common such algebraic identities below leaving for the reader some of the routine proofs.

PROPOSITION 1. *Let* $\mathbb{F}$ *be a set on which there are defined binary operations of addition and multiplication.*

(1) *The addition axioms imply*
  (a) $x+y = x+z \Longrightarrow y = z$*,*
  (b) $x+y = x \Longrightarrow y = 0$*,*
  (c) $x+y = 0 \Longrightarrow y = -x$*,*
  (d) $-(-x) = x$*.*
(2) *The multiplication axioms imply*

(a) $x \neq 0$ *and* $xy = xz \Longrightarrow y = z$,
(b) $x \neq 0$ *and* $xy = x \Longrightarrow y = 1$,
(c) $x \neq 0$ *and* $xy = 1 \Longrightarrow y = \frac{1}{x}$,
(d) $x \neq 0 \Longrightarrow \frac{1}{\frac{1}{x}} = x$.

(3) *The field axioms imply*
(a) $0x = 0$,
(b) $x \neq 0$ *and* $y \neq 0 \Longrightarrow xy \neq 0$,
(c) $(-x)\,y = -(xy) = x\,(-y)$,
(d) $(-x)\,(-y) = xy$.

By way of illustration we prove the final equality $(-x)\,(-y) = xy$ by a method that also establishes $(1)\,(a)\,(c)\,(d)$ and $(3)\,(a)\,(c)$ along the way (much shorter proofs also exist). For this we begin with the additive cancellation property $(1)\,(a)$: if $x + y = x + z$ then

$$\begin{aligned} y &= 0 + y = (-x + x) + y = -x + (x + y) \\ &= -x + (x + z) \qquad \text{by assumption} \\ &= (-x + x) + z = 0 + z = z. \end{aligned}$$

Taking $z = -x$ this gives $(1)\,(c)$ (uniqueness of additive inverses), and since $(-x) + x = 0$, $(1)\,(c)$ then gives $x = -(-x)$, which is $(1)\,(d)$. Next we note that

$$(-x)\,y + xy = (-x + x)\,y = 0y = 0, \tag{3.2}$$

where the final equality follows from applying additive cancellation $(1)\,(a)$ to

$$0y + 0y = (0 + 0)\,y = 0y = 0y + 0.$$

By applying $(1)\,(c)$ to (3.2) we obtain

$$xy = -((-x)\,y)\,. \tag{3.3}$$

If we interchange x and y in (3.3) and use multiplicative commutativity, we also obtain

$$xy = yx = -((-y)\,x) = -(x\,(-y))\,. \tag{3.4}$$

Finally, with x replaced by $-x$ and y replaced by $-y$ in (3.3) we have

$$(-x)\,(-y) = -((-(-x))\,(-y)) = -(x\,(-y))\,,$$

which when combined with (3.4) yields $(-x)\,(-y) = xy$ as required.

Now we combine the field and order properties. By $x > y$ we mean $y < x$.

Definition 3. *An* ordered field *is a field* $\mathbb{F}$ *together with an order* $<$ *on the set* $\mathbb{F}$ *where the field and order structures are connected by the following two additional axioms:*
(1) $x + y < x + z$ *if* $x, y, z \in \mathbb{F}$ *and* $y < z$,
(2) $xy > 0$ *if* $x, y \in \mathbb{F}$ *and both* $x > 0$ *and* $y > 0$.

Example 4. *The field of rational numbers* $\mathbb{Q}$ *is an ordered field with the usual order, but for* p *a prime, there is* no *order on the field* $\mathbb{F}_p$ *that satisfies Definition 3.*

All of the customary rules for manipulating inequalities in the rational numbers hold also in any ordered field. We state the most common such properties below, without giving the routine proofs.

PROPOSITION 2. *The following hold in any ordered field.*

(1) $x > 0$ *if and only if* $-x < 0$,
(2) $xy < xz$ *if* $x > 0$ *and* $y < z$,
(3) $xy > xz$ *if* $x < 0$ *and* $y < z$,
(4) $x^2 > 0$ *if* $x \neq 0$,
(5) $1 > 0$,
(6) $0 < \frac{1}{y} < \frac{1}{x}$ *if* $0 < x < y$.

Now we come to the most important property an ordered field can have, one that is essential for the success of analysis, but is *not* satisfied in the ordered field of rational numbers $\mathbb{Q}$.

DEFINITION 4. *Let* $<$ *be an order on a set* S.

(1) *We say that* $x \in S$ *is an* upper bound *for a subset* E *of* S *if*
$$y \leq x \text{ for all } y \in E.$$
(2) *We say that a subset* E *is* bounded above *if it has at least one upper bound.*
(3) *We say that* $x \in S$ *is the* least upper bound *for a subset* E *of* S *if* x *is an upper bound for* E *and if* z *is any other upper bound for* E, *then* $x \leq z$. *In this case we write*
$$x = \sup E.$$

Clearly the least upper bound of a subset E, if it exists, is unique. Consider the ordered set of rational numbers $\mathbb{Q}$. Then 3 is an upper bound for the interval $E = [0, 3] = \{x \in \mathbb{Q} : 0 \leq x \leq 3\}$, and so are π, 4 and 2^{100}. In fact it is easy to see that 3 is the *least* upper bound for $[0, 3]$. An example of a subset that has no least upper bound is the semiinfinite interval $[0, \infty) = \{x \in \mathbb{Q} : 0 \leq x < \infty\}$, since it has no upper bounds at all! A more substantial example of a *bounded* set that has no least upper bound is the set A defined in (2.1).

There are corresponding definitions of *lower bound*, *bounded below*, *greatest lower bound* and $\inf E$, whose formulations we leave to the reader.

DEFINITION 5. *An ordered set* S *has the* Least Upper Bound Property *if every subset* E *of* S *that is bounded above has a least upper bound.*

The ordered set of rational numbers $\mathbb{Q}$ *fails* to have this crucial property, as evidenced by the existence of the set A in (2.1). An example of a nontrivial ordered set with the Least Upper Bound Property is the set of all ordinal numbers equal to or less than the first uncountable ordinal.

REMARK 4. *If* S *has the Least Upper Bound Property, it also has the* Greatest Lower Bound Property*: every subset* E *of* S *that is bounded below has a greatest lower bound. To see this, suppose* E *is bounded below and let* L *be the nonempty set of lower bounds. Then* L *is bounded above by every element of* E *and in particular* $\alpha = \sup L$ *exists. Now* $\alpha = \inf E$ *follows from the following two facts:*

(1) $\alpha \in L$ *since if* $\gamma < \alpha$, *then* γ *cannot be an upper bound of* L, *hence* $\gamma \notin E$ *since every element of* E *is an upper bound of* L. *Thus* $\alpha \leq x$ *for every* $x \in E$ *and so* $\alpha \in L$.
(2) $\beta \notin L$ *if* $\beta > \alpha$ *since* α *is an upper bound of* L.

It turns out that the only ordered field that has the Least Upper Bound Property is (up to isomorphism) the ordered field of real numbers $\mathbb{R}$, which we have not yet constructed. Before embarking on the construction of the real numbers using Dedekind cuts, it will be useful to derive some consequences of the Least Upper Bound Property in an ordered field. Just so we can be certain we are not working in a vaccuum, we state the basic existence theorem whose proof is deferred to the end of this chapter.

THEOREM 1. *There exists an ordered field $\mathbb{R}$ having the Least Upper Bound Property. Moreover, such a field is uniquely determined up to isomorphism (of ordered fields) and contains (an isomorphic copy of) the rational field $\mathbb{Q}$ as a subfield.*

Assuming this existence theorem for the moment we derive some properties of ordered fields with the Least Upper Bound Property. We note that we could also prove these properties by appealing to the explicit construction of the real numbers by Dedekind cuts below, but the approach used here is more streamlined in that it avoids the complexities inherent in the construction of the reals. We begin with two familiar properties shared by the field of rational numbers.

PROPOSITION 3. *Let $x, y \in \mathbb{R}$.*

(1) *(Archimedian property) If $x > 0$, then there is a positive integer n such that $nx > y$.*
(2) *(density of rationals) If $x < y$ then there is $p \in \mathbb{Q}$ such that $x < p < y$.*

Proof: To prove assertion (1) by contradiction, let $E = \{nx : n \in \mathbb{N}\}$. If (1) were false, then y would be an upper bound for E and consequently $\alpha = \sup E$ would exist. Since $x > 0$, we would have $\alpha - x < \alpha$ and thus that $\alpha - x$ could not be an upper bound for E. But then there would be some nx greater than $\alpha - x$ and this gives

$$\begin{aligned} \alpha &= (\alpha - x) + x \\ &< nx + x \\ &= (n+1)\,x \in E, \end{aligned}$$

which contradicts the assumption that α is an upper bound for E.

To prove assertion (2), use assertion (1) to choose $n \in \mathbb{N}$ such that $n\,(y - x) > 1$. Use assertion (1) twice more to obtain integers m_1 and m_2 satisfying $m_1 > nx$ and $m_2 > -nx$. Thus we have both

$$n\,(y - x) > 1 \text{ and } -m_2 < nx < m_1.$$

Because $m_1 - (-m_2) > nx + (-nx) = 0$, i.e. $m_1 - (-m_2) \geq 1$, it follows that there is an integer m lying between $-m_2$ and m_1 such that

$$m - 1 \leq nx < m.$$

Combining inequalities yields

$$nx < m \leq 1 + nx < ny,$$

and since $n > 0$ we obtain

$$x < \frac{m}{n} < y.$$

Similar reasoning can be used to obtain the existence of positive n^{th} roots of positive numbers in an ordered field with the least upper bound property. This property is *not* shared by the field of rational numbers.

PROPOSITION 4. *(existence of n^{th} roots) If x is a positive real number and n is a positive integer, then there exists a unique positive real number y satisfying $y^n = x$.*

Sketch of the proof: Let $E = \{z \in \mathbb{R} : 0 < z \text{ and } z^n < x\}$. One can show that E is nonempty and bounded above, hence $y = \sup E$ exists. Using an argument similar to that following (2.1) one can now show that each of the inequalities $y^n < x$ and $y^n > x$ leads to a contradiction, leaving only the possibility that $y^n = x$. For details of these arguments see page 10 of [**3**].

Note that $\sup A = \sqrt{2}$ where A is the set in (2.1).

COROLLARY 1. *If x and y are positive real numbers and n is a positive integer, then $x^{\frac{1}{n}} y^{\frac{1}{n}} = (xy)^{\frac{1}{n}}$.*

Proof: By the commutativity of multiplication we have

$$\begin{aligned}\left(x^{\frac{1}{n}} y^{\frac{1}{n}}\right)^n &= \left(x^{\frac{1}{n}} y^{\frac{1}{n}}\right)\left(x^{\frac{1}{n}} y^{\frac{1}{n}}\right)...\left(x^{\frac{1}{n}} y^{\frac{1}{n}}\right) \\ &= \left(x^{\frac{1}{n}}\right)\left(x^{\frac{1}{n}}\right)...\left(x^{\frac{1}{n}}\right) \times \left(y^{\frac{1}{n}}\right)\left(y^{\frac{1}{n}}\right)...\left(y^{\frac{1}{n}}\right) \\ &= \left(x^{\frac{1}{n}}\right)^n \left(y^{\frac{1}{n}}\right)^n = xy.\end{aligned}$$

By the uniqueness assertion of Proposition 4 we then conclude that $x^{\frac{1}{n}} y^{\frac{1}{n}} = (xy)^{\frac{1}{n}}$.

4. The complex field

Property (4) of Proposition 2 on ordered fields shows that there is *no* real number x satisfying the equation $x^2 = -1$. To remedy this situation, we define the complex field $\mathbb{C}$ to be the field obtained from the real field $\mathbb{R}$ by adjoining an abstract symbol i that is declared to satisfy the equation

$$i^2 = -1. \tag{4.1}$$

Thus $\mathbb{C}$ consists of all expressions of the form

$$z = x + iy, \qquad x, y \in \mathbb{R},$$

which can be identified with the "points in the plane" by associating $z = x + iy \in \mathbb{C}$ with $(x, y) \in \mathbb{R} \times \mathbb{R}$ in the plane. The field structure on $\mathbb{C}$ uses the multiplication rule derived from (4.1) by

$$\begin{aligned} zw &= (x + iy)(u + iv) \\ &= \left(xu + i^2 yv\right) + i(xv + yu) \\ &= (xu - yv) + i(xv + yu), \end{aligned} \tag{4.2}$$

where $z = x + iy$ and $w = u + iv$. For the most part, straightforward calculations show that this multiplication and the usual addition derived from vectors in the plane $\mathbb{R} \times \mathbb{R}$,

$$(x + iy) + (u + iv) = (x + u) + i(y + v),$$

satisfy the addition axioms, the multiplication axioms and the distributive law of a field. Only the existence of a multiplicative inverse needs some elaboration. For this we define

DEFINITION 6. *Suppose $z = x + iy \in \mathbb{C}$. The complex conjugate $\overline{z}$ ofz is defined to be*

$$\overline{z} = x - iy.$$

Now

$$z\overline{z} = (x + iy)(x - iy) = x^2 - (iy)^2 + i\{yx - xy\} = x^2 + y^2,$$

and by Proposition 4, the nonnegative real number $\sqrt{x^2 + y^2}$ exists and is unique. By Pythagoras' theorem,

$$\sqrt{x^2 + y^2} = \sqrt{z\overline{z}}$$

is the distance between the complex numbers 0 and z when they are viewed as the points $(0, 0)$ and (x, y) in the plane. We define

$$|z| = \sqrt{z\overline{z}}, \qquad z \in \mathbb{C},$$

called the absolute value of z, and note that for $z \in \mathbb{C} \setminus \{0\}$, the multiplicative inverse of z is given by $z^{-1} = \frac{\overline{z}}{|z|^2}$ since

$$z\left(z^{-1}\right) = z\frac{\overline{z}}{|z|^2} = \frac{z\overline{z}}{|z|^2} = \frac{|z|^2}{|z|^2} = 1.$$

We now make three observations.

(1) An immediate consequence of property (4) of Proposition 2 is that there is *no* order on $\mathbb{C}$ that makes it into an ordered field with this field structure.
(2) It is a fundamental theorem in algebra, in fact it is called *the* fundamental theorem of algebra, that we do not need to adjoin any further solutions of polynomial equations: *every* polynomial equation

$$z^n + a_{n-1}z^{n-1} + ... + a_1 z + a_0 = 0$$

has a solution z in the complex field $\mathbb{C}$. Here the coefficients $a_0, a_1, ..., a_{n-1}$ are complex numbers.
(3) If we associate $z = x + iy$ to the matrix $\begin{bmatrix} x & -y \\ y & x \end{bmatrix}$, then this multiplication corresponds to matrix multiplication:

$$\begin{aligned} [z]\,[w] &= \begin{bmatrix} x & -y \\ y & x \end{bmatrix}\begin{bmatrix} u & -v \\ v & u \end{bmatrix} \\ &= \begin{bmatrix} xu - yv & -xv - yu \\ yu + xv & -yv + xu \end{bmatrix} = [zw]\,. \end{aligned} \tag{4.3}$$

Since the matrix

$$\begin{bmatrix} x & -y \\ y & x \end{bmatrix} = r\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

is dilation by the nonnegative number $r = \sqrt{x^2 + y^2} = |z|$ and rotation by the angle $\theta = \tan^{-1}\frac{y}{x}$ in the counterclockwise direction, we see that if z has polar coordinates (r, θ) and w has polar coordinates (s, ϕ), then zw

has polar coordinates $(rs, \theta + \phi)$. Finally we note that the inverse of the matrix $M = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}$ is given by

$$M^{-1} = \frac{1}{\det M} [coM]^t = \frac{1}{x^2 + y^2} \begin{bmatrix} x & y \\ -y & x \end{bmatrix} = \begin{bmatrix} \frac{x}{x^2+y^2} & \frac{y}{x^2+y^2} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix},$$

which agrees with $z^{-1} = \frac{\overline{z}}{|z|^2} = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2}$ (M is the matrix representation of the real linear map induced on $\mathbb{R}^2$ by the map of complex multiplication on $\mathbb{C} = \mathbb{R}^2$ by $z = x + iy$).

Finally we give some simple properties of the complex conjugate and absolute value functions. If $z = x + iy$ we write $\operatorname{Re} z = x$ and $\operatorname{Im} z = y$.

Proposition 5. *Let $\overline{z}$ and $|z|$ denote the complex conjugate and absolute value of z.*

(1) *Suppose $z, w \in \mathbb{C}$. Then*
  (a) $\overline{z+w} = \overline{z} + \overline{w}$, $\overline{(zw)} = (\overline{z})(\overline{w})$ *and* $z + \overline{z} = 2\operatorname{Re} z$,
  (b) $|0| = 0$ *and* $|z| > 0$ *unless* $z = 0$,
  (c) $|\overline{z}| = |z|$,
  (d) $|zw| = |z|\,|w|$,
  (e) $|\operatorname{Re} z| \leq |z|$,
  (f) $|z + w| \leq |z| + |w|$.

(2) *(Cauchy-Schwarz inequality) Suppose $z_1, ..., z_n \in \mathbb{C}$ and $w_1, ..., w_n \in \mathbb{C}$. Then*

$$\left| \sum_{j=1}^{n} z_j \overline{w_j} \right|^2 \equiv |z_1\overline{w_1} + ... + z_n\overline{w_n}|^2 \leq \left(\sum_{j=1}^{n} |z_j|^2 \right) \left(\sum_{j=1}^{n} |w_j|^2 \right).$$

Proof: Assertions (1) (a) (b) (c) (e) are easy. If $z = x + iy$ and $w = u + iv$ then from (4.2),

$$\begin{aligned} |zw|^2 &= |(xu - yv) + i(xv + yu)|^2 \\ &= (xu - yv)^2 + (xv + yu)^2 \\ &= x^2u^2 - 2xuyv + y^2v^2 + x^2v^2 + 2xvyu + y^2u^2 \\ &= (x^2 + y^2)(u^2 + v^2) = |z|^2 |w|^2, \end{aligned}$$

and now the uniqueness assertion of Proposition 4 proves (1) (d).

Next we compute

$$\begin{aligned} |z + w|^2 &= (z + w)(\overline{z + w}) = (z + w)(\overline{z} + \overline{w}) \\ &= z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} \\ &= |z|^2 + 2\operatorname{Re}(z\overline{w}) + |w|^2 \\ &\leq |z|^2 + 2|z\overline{w}| + |w|^2 \\ &= |z|^2 + 2|z|\,|w| + |w|^2 = (|z| + |w|)^2, \end{aligned}$$

and the uniqueness assertion of Proposition 4 now proves (1) (f).

Finally, to obtain (2), set

$$Z = \sum_{j=1}^{n} |z_j|^2 \text{ and } W = \sum_{j=1}^{n} |w_j|^2 \text{ and } D = \sum_{j=1}^{n} z_j \overline{w_j},$$

so that we must prove

$$|D|^2 \le ZW. \tag{4.4}$$

If $W = 0$ then both sides of (4.4) vanish. Otherwise, we have

$$\begin{aligned}
\sum_{j=1}^{n} |Wz_j - Dw_j|^2 &= \sum_{j=1}^{n} (Wz_j - Dw_j)\left(W\overline{z_j} - \overline{D}\overline{w_j}\right) \\
&= W^2 \sum_{j=1}^{n} |z_j|^2 - W\overline{D} \sum_{j=1}^{n} z_j \overline{w_j} - DW \sum_{j=1}^{n} w_j \overline{z_j} + |D|^2 \sum_{j=1}^{n} |w_j|^2 \\
&= W^2 Z - W\overline{D}D - DW\overline{D} - |D|^2 W \\
&= W^2 Z - W|D|^2 = W\left(WZ - |D|^2\right),
\end{aligned}$$

and since $W > 0$ we obtain

$$WZ - |D|^2 = \frac{1}{W} \sum_{j=1}^{n} |Wz_j - Dw_j|^2 \ge 0.$$

4.1. Euclidean spaces. For $\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbb{R} \times \mathbb{R} \times ... \times \mathbb{R} \equiv \mathbb{R}^n$, we define

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2},$$

and interpret $\|\mathbf{x}\|$ as the distance from the point $\mathbf{x}$ to the origin $0 = (0, 0, ..., 0)$, which is reasonable since it agrees with Pythagoras' theorem. We call $\mathbb{R}^n$ the Euclidean space of dimension n. For $\mathbf{z}, \mathbf{w} \in \mathbb{R}^n$, we define the *dot product* of $\mathbf{z}$ and $\mathbf{w}$ by

$$\mathbf{z} \cdot \mathbf{w} = z_1 w_1 + z_2 w_2 + ... + z_n w_n = \sum_{j=1}^{n} z_j w_j.$$

The Cauchy-Schwarz inequality, when restricted to real numbers, says that

$$|\mathbf{z} \cdot \mathbf{w}| \le \|\mathbf{z}\| \, \|\mathbf{w}\|, \qquad \mathbf{z}, \mathbf{w} \in \mathbb{R}^n.$$

Remark 5. *The proof of the Cauchy-Schwarz inequality given above is motivated by the fact that in a Euclidean space, the point on the line through* 0 *and* $\mathbf{w}$ *that is closest to* $\mathbf{z}$ *is the projection* $P\mathbf{z}$ *of* $\mathbf{z}$ *onto the line through* 0 *and* $\mathbf{w}$ *given by*

$$P\mathbf{z} = \left(\mathbf{z} \cdot \frac{\mathbf{w}}{\|\mathbf{w}\|}\right) \frac{\mathbf{w}}{\|\mathbf{w}\|} = \frac{\mathbf{z} \cdot \mathbf{w}}{\|\mathbf{w}\|^2} \mathbf{w}.$$

Then

$$\begin{aligned}
\|\mathbf{z} - P\mathbf{z}\|^2 &= \sum_{j=1}^{n} \left| z_j - \frac{z \cdot w}{\|w\|^2} w_j \right|^2 \\
&= \frac{1}{\|\mathbf{w}\|^4} \sum_{j=1}^{n} \left| \|\mathbf{w}\|^2 z_j - (\mathbf{z} \cdot \mathbf{w}) w_j \right|^2 \\
&= \frac{1}{\|\mathbf{w}\|^4} \sum_{j=1}^{n} |Wz_j - Dw_j|^2 .
\end{aligned}$$

5. Dedekind's construction of the real numbers

Recall that a Dedekind cut α is a subset of $\mathbb{Q}$ satisfying (3.1),

$$\begin{aligned} \alpha &\neq \emptyset \text{ and } \alpha \neq \mathbb{Q}, \\ p &\in \alpha \text{ and } q < p \text{ implies } q \in \alpha, \\ p &\in \alpha \text{ implies there is } q \in \alpha \text{ with } p < q. \end{aligned}$$

We set

$$\mathbb{R} = \{\alpha : \alpha \text{ is a cut}\},$$

and define an order $<$ and two binary operations, addition $+$ and multiplication $\cdot$, on the set $\mathbb{R}$ and then demonstrate that $\mathbb{R}$ satisfies the axioms for an ordered field with the Least Upper Bound Property. We proceed in six steps, giving proofs only when there is some trick involved, or the result is especially important. The letters p, q, r, s, t always denote rational numbers and the Greek letters $\alpha, \beta, \gamma, \theta$ always denote cuts. See pages 17-21 of [**3**] for the details.

Step 1: Define $\alpha < \beta$ if α is a proper subset of β. Then $(\mathbb{R}, <)$ is an ordered set.

Step 2: $(\mathbb{R}, <)$ has the Least Upper Bound Property.

Proof: To see this, suppose that E is a nonempty subset of $\mathbb{R}$ that is bounded above by $\beta \in \mathbb{R}$. Define

$$\gamma = \bigcup_{\alpha \in E} \alpha.$$

One can now show that γ is a cut ($\gamma \neq \emptyset$ since there exists $\alpha\,(\neq \emptyset) \in E$ and then $\alpha \subset \gamma$; $\gamma \neq \mathbb{Q}$ since $\gamma \subset \beta$ and $\beta \neq \mathbb{Q}$; if $p \in \gamma$, then there is $\alpha \in E$ with $p \in \alpha$, and it follows that every q less than p is in $\alpha \subset \gamma$ and there is r in $\alpha \subset \gamma$ that is larger than p), and clearly γ is then an upper bound for E since $\alpha \subset \gamma$ for all $\alpha \in E$. Moreover, γ is the *least* upper bound, written $\gamma = \sup E$, since any upper bound must contain at least each set $\alpha \in E$. Note how easily we obtained the Least Upper Bound Property by this construction!

Step 3: If $\alpha, \beta \in \mathbb{R}$, define

$$\alpha + \beta = \{p + q : p \in \alpha \text{ and } q \in \beta\}.$$

Also set

$$\theta = \{p \in \mathbb{Q} : p < 0\}.$$

Then $\alpha + \beta$ and θ are cuts and using θ as the additive identity 0, the Addition Axioms for a field hold. In fact more is true: if α is a cut and β is *any* nonempty set that is bounded above, then $\alpha + \beta$ is a cut.

Proof: If $p = r + s \in \alpha + \beta$ and $q < p$, then $q = (q - p + r) + s \in \alpha + \beta$ since $q - p + r < r$ and α is a cut. Furthermore, there is $t \in \alpha$ with $t > r$ and so $t + s \in \alpha + \beta$ with $t + s > r + s = p$. Obviously θ is a cut. Next, $\alpha + \theta \subset \alpha$ and if $p \in \alpha$, then there is $r \in \alpha$ with $r > p$ and so $p = r + (p - r) \in \alpha + \theta$, and this shows that $\alpha + \theta = \alpha$ for all $\alpha \in \mathbb{R}$. It requires only a bit more effort to show that the inverse of $\alpha \in \mathbb{R}$ is given by the set

$$-\alpha \equiv \{p \in \mathbb{Q} : \text{there exists } r > 0 \text{ such that } -p - r \notin \alpha\}.$$

Indeed, it is not too hard to show that $-\alpha$ is a cut. To see the more delicate fact that

$$\alpha + (-\alpha) = \theta, \tag{5.1}$$

we first note that $\alpha + (-\alpha) \subset \theta$ since if $q \in \alpha$ and $r \in -\alpha$, then $-r \notin \alpha$, hence $q < -r$, hence $q + r < 0$. Conversely, pick $s \in \theta$ and set $t = -\frac{s}{2} > 0$. By the Archimedian property of the rational numbers $\mathbb{Q}$, there is $n \in \mathbb{N}$ such that

$$nt \in \alpha \text{ but } (n+1)\,t \notin \alpha.$$

Set $p = -(n+2)\,t$.

Remark 6. *It is helpful at this point to suppose that α corresponds to a point on the line to the right of 0, and to draw the players in the proof from left to right on the line:*

$$p < -(n+1)\,t < -\alpha < -nt < -t < 0 < t < nt < \alpha < (n+1)\,t < -p.$$

Now $p \in -\alpha$ since $-p - t = (n+1)\,t \notin \alpha$. Since $nt \in \alpha$ we thus have

$$s = -2t = nt + p \in \alpha + (-\alpha).$$

This proves that $\theta \subset \alpha + (-\alpha)$ and completes the proof of (5.1).

Step 4: If $\alpha, \beta, \gamma \in \mathbb{R}$ and $\beta < \gamma$, then $\alpha + \beta < \alpha + \gamma$.

Proof: This is easy to prove using the cancellation law for addition in Proposition 1 (1) (a). Indeed, when cuts are considered as subsets of rational numbers, we clearly have $\alpha + \beta \subset \alpha + \gamma$. If we had equality $\alpha + \beta = \alpha + \gamma$, then Proposition 1 (1) (a) shows that $\beta = \gamma$, a contradiction. Note that Proposition 1 (1) applies here since we have shown in Step 3 that the addition axioms hold.

Step 5: If $\alpha, \beta > \theta$, define

$$\begin{aligned} \alpha \cdot \beta \;\; &= \;\; \{p \in \mathbb{Q} : p \leq qr \text{ for some choice of} \\ &\qquad q \in \alpha \text{ with } q > 0 \text{ and } r \in \beta \text{ with } r > 0\}. \end{aligned}$$

For general $\alpha, \beta \in \mathbb{R}$, define $\alpha \cdot \beta$ appropriately. Then $(\mathbb{R}, <, +, \cdot)$ is an ordered field with the Least Upper Bound Property.

Proof: The proof of the multiplication axioms is somewhat bothersome due to the different definitions of product $\alpha \cdot \beta$ according to the signs of α and β. We omit the remaining tedious details in the proof of Step 5.

Step 6: To each $q \in \mathbb{Q}$ we associate the set

$$\Gamma(q) \equiv \{p \in \mathbb{Q} : p < q\}.$$

Then $\Gamma(q)$ is a cut and

$$\begin{aligned} \Gamma(r+s) &= \Gamma(r) + \Gamma(s), \\ \Gamma(rs) &= \Gamma(r) \cdot \Gamma(s), \\ \Gamma(r) &< \Gamma(s) \Longleftrightarrow r < s. \end{aligned}$$

Thus the map $\Gamma : \mathbb{Q} \to \mathbb{R}$ is an ordered field isomorphism from the rational numbers $\mathbb{Q}$ into the real numbers $\mathbb{R}$, and this is the sense in which we mean that the real numbers $\mathbb{R}$ contain a copy of the rational numbers $\mathbb{Q}$.

Remark 7. *One might reasonably ask why in the definition of cut (3.1) we had to include the third condition requiring the cut to have no largest element:*

$$p \in \alpha \textit{ implies there is } q \in \alpha \textit{ with } p < q.$$

However, without this condition, there are additional cuts, namely those with a largest rational element:

$$r^* \equiv \{p \in \mathbb{Q} : p \leq r\}, \qquad \textit{for } r \in \mathbb{Q}.$$

We refer to these additional cuts as closed cuts, and to the original cuts as open cuts. A cut that is either closed or open is said to be a generalized cut. Suppose we extend the definition of addition to generalized cuts in the standard way by taking all possible sums of pairs, one element from each cut. The key property to observe then is that $\alpha + \beta$ *is an* open *cut provided at least one of* α *and* β *is open (see Step 3 above). Thus the usual zero element* 0 *can no longer serve as the additive identity for the set of generalized cuts. It is not hard to see however that the closed cut*

$$0^* \equiv \{p \in \mathbb{Q} : p \leq 0\}$$

has the required additive identity property $0^* + \alpha = \alpha$ *for all generalized cuts* α *- in fact* 0^* *is the* only *generalized cut with this property. Now comes the problem. An* open *cut* α *cannot have an additive inverse since the result of adding* any *generalized cut to* α *must also be open - and in particular cannot equal the* closed *cut* 0^**.*

CHAPTER 2

Cardinality of sets

Dedekind was the first to define an infinite set as one to which the paradoxes of Galileo and Bolzano applied (there are as many perfect squares as there are integers; there are as many even integers as there are integers; and there are as many points in the interval $[0,1]$ as there are in $[0,2]$):

- A system S is said to be *infinite* if it is similar to a proper part of itself; in the contrary case S is said to be a *finite* system.

In other words, a set S was defined to be infinite by Dedekind if there existed a one-to-one correspondence between S and a *proper* subset of itself. However, Dedekind's definition gave no hint that there might be different 'sizes' of infinity, and the creation of this revolutionary concept had to await the imagination of Georg Cantor.

Definition 7. *Two sets A and B are said to have the* same cardinality *or are said to be* equivalent*, written $A \sim B$, if there is a one-to-one onto map $\varphi : A \to B$. Let $n \in \mathbb{N}$. A set E is said to have cardinality n if it is equivalent to the set*

$$J_n \equiv \{1, 2, 3, ..., n-1, n\},$$

in which case it is said to be finite. A set E is said to be countable *if it is equivalent to the set of natural numbers $\mathbb{N}$. If a set is neither finite nor countable, it is said to be* uncountable*.*

The relation $\sim$ of having the same cardinality is an equivalence relation, meaning that it satisfies

(1) (reflexivity) $A \sim A$,
(2) (symmetry) $A \sim B \Longrightarrow B \sim A$,
(3) (transitiviy) $A \sim B$ and $B \sim C \Longrightarrow A \sim C$.

These equivalence classes are called *cardinal numbers* since they measure the size of sets up to bijections. Cantor showed at least two surprising results regarding cardinality: first, that the set of rational numbers is countable and second, that the set of real numbers is uncountable. Both demonstrations involved a notion of diagonalization.

To show that the rational numbers $\mathbb{Q}$ are countable, Cantor arranged the positive rational numbers $\mathbb{Q}_+$ in an infinite matrix $\left[\frac{m}{n}\right]_{m,n=1}^{\infty}$;

$$\left[\begin{array}{ccccccccc}
\frac{1}{1} & & \frac{1}{2} & & \frac{1}{3} & & \frac{1}{4} & & \cdots \\
& \nearrow & & \nearrow & & \nearrow & & \nearrow & \\
\frac{2}{1} & & \frac{2}{2} & & \frac{2}{3} & & \frac{2}{4} & & \cdots \\
& \nearrow & & \nearrow & & \nearrow & & & \\
\frac{3}{1} & & \frac{3}{2} & & \frac{3}{3} & & \frac{3}{4} & & \cdots \\
& \nearrow & & \nearrow & & & & & \\
\frac{4}{1} & & \frac{4}{2} & & \frac{4}{3} & & \frac{4}{4} & & \cdots \\
\vdots & \nearrow & \vdots & & \vdots & & \vdots & & \ddots
\end{array}\right],$$

and then defined a map $s : \mathbb{N} \to \mathbb{Q}_+$ by following the upward sloping diagonals in succession, taking only those fractions that have not yet appeared:

$$\begin{array}{rcl}
s(1) & = & \frac{1}{1}; \\
s(2) & = & \frac{2}{1}; \quad s(3) = \frac{1}{2}; \\
s(4) & = & \frac{3}{1}; \quad s(5) = \frac{1}{3} \quad (\frac{2}{2} = s(1) \text{ was skipped}); \\
s(6) & = & \frac{4}{1}; \quad s(7) = \frac{3}{2}; \quad s(8) = \frac{2}{3}; \quad s(9) = \frac{1}{4}; \\
s(10) & = & \frac{5}{1}; \quad s(11) = \frac{1}{5} \quad (\frac{4}{2} = s(2), \frac{3}{3} = s(1), \frac{2}{4} = s(3) \text{ were all skipped}); \\
& & \vdots
\end{array}$$

Clearly the map s is one-to-one and onto, thus demonstrating that $\mathbb{N} \sim \mathbb{Q}_+$. It is now a simple matter to use s to construct a one-to-one onto map $t : \mathbb{N} \to \mathbb{Q}$ (**exercise**: do this!) that shows $\mathbb{N} \sim \mathbb{Q}$.

To show that the real numbers are uncountable, we begin with a famous paradox of Russell. Define a set S by the rule

$$a \in S \Leftrightarrow a \notin a,$$

i.e. S consists of all sets a that are *not* members of themselves. Then we have the following paradox:

- If $S \in S$, then by the very definition of S it must be the case that $S \notin S$, a contradiction.
- On the other hand if $S \notin S$, then by the very definition of S it must be the case that $S \in S$, again a contradiction.

One way out of this paradox is to note that we have *never* seen a set a that is a member of itself. Thus we expect that S is actually the collection of *all* sets. If we simply *disallow* the collection of all sets as a set, Russell's paradox dissolves. This type of thinking eventually led to the Zermelo-Frankel set theory in use today.

Russell's paradox suggests the following proof that the power set

$$\mathcal{P}(\mathbb{N}) \equiv \{E : E \subset \mathbb{N}\}$$

of the natural numbers, i.e. the set of all subsets of $\mathbb{N}$, is uncountable. Indeed, assume in order to derive a contradiction, that $\mathcal{P}(\mathbb{N})$ is countable. Then we can list all the elements of $\mathcal{P}(\mathbb{N}) = \{E^m\}_{m=1}^{\infty}$ in a vertical column:

$$\begin{bmatrix} E^1 \\ E^2 \\ E^3 \\ \vdots \end{bmatrix}.$$

Now each subset E^m is uniquely determined by its characteristic function, i.e. the sequence $\{s_n^m\}_{n=1}^{\infty} = \{s_1^m, s_2^m, s_3^m, ...\}$ of 0's and 1's defined by

$$s_n^m = \begin{cases} 0 & \text{if} \quad n \notin E_m \\ 1 & \text{if} \quad n \in E_m \end{cases}.$$

Replace each subset E^m in the vertical column by the infinite row of 0's and 1's determined by $\{s_n^m\}_{n=1}^{\infty}$ to get an infinite matrix of 0's and 1's:

$$\begin{bmatrix} s_1^1 & s_2^1 & s_3^1 & \cdots \\ s_1^2 & s_2^2 & s_3^2 & \\ s_1^3 & s_2^3 & s_3^3 & \\ \vdots & & & \ddots \end{bmatrix}.$$

Now consider the *anti-diagonal* or *Russell* sequence $\{r_n\}_{n=1}^{\infty}$ given by

$$r_n = 1 - s_n^n. \tag{0.2}$$

This is a sequence of 0's and 1's that is *not* included in the list

$$\begin{bmatrix} \{s_n^1\}_{n=1}^{\infty} \\ \{s_n^2\}_{n=1}^{\infty} \\ \{s_n^3\}_{n=1}^{\infty} \\ \vdots \end{bmatrix},$$

since for each m, the sequences $\{s_n^m\}_{n=1}^{\infty}$ and $\{r_n\}_{n=1}^{\infty}$ differ in the m^{th} entry: $s_m^m \neq r_m$ by (0.2). Thus the set $E = \{n : r_n = 1\}$ whose characteristic function is the sequence $\{r_n\}_{n=1}^{\infty}$ satisfies

$$n \in E \Leftrightarrow r_n = 1 \Leftrightarrow s_n^n = 0 \Leftrightarrow n \notin E_n,$$

and hence is the set of n such that n *is not a member of* E_n (reminiscent of Russell's paradox). It follows that E is *not* included in the list $\{E^m\}_{m=1}^{\infty}$. This contradiction shows that the power set $\mathcal{P}(\mathbb{N})$ is uncountable. Equivalently, this shows that the set of all sequences consisting of 0's and 1's is uncountable.

To see from this that the real numbers are uncountable, express each real number s in the interval $(0, 1]$ as a binary fraction

$$s = \frac{s_1}{2} + \frac{s_2}{2^2} + ... + \frac{s_n}{2^n} + ... \equiv 0.s_1 s_2 ... s_n ...$$

where the sequence $\{s_n\}_{n=1}^{\infty}$ does not end in an infinite string of 0's. Since the set of such fractions is uncountable (in fact its equivalence with $\mathcal{P}(\mathbb{N})$ follows from the argument above with just a little extra work), we conclude that the interval $(0, 1]$ is uncountable, and then so is $\mathbb{R}$. We will return to this argument later.

We now turn to the task of making the previous arguments more rigorous. We begin with a careful definition of 'sequence'.

DEFINITION 8. *A* sequence *is a function f defined on the natural numbers $\mathbb{N}$. If $f(n) = s_n$ for all $n \in \mathbb{N}$, the values s_n are called the* terms *of the sequence, and we often denote the sequence f by $\{s_n\}_{n=1}^{\infty}$ or even $\{s_1, s_2, s_3, ...\}$.*

Thus we may regard a countable set as the range of a sequence of *distinct* terms, and in fact we used this point of view when we assumed above that $\mathcal{P}(\mathbb{N})$ was countable and then listed the elements of $\mathcal{P}(\mathbb{N})$ in a vertical column. The next lemma proves the intuitive fact that *'countable is the smallest infinity'*.

LEMMA 1. *Every infinite subset of a countable set is countable.*

Proof: Suppose A is countable and E is an infinite subset of A. Represent A as the range of a sequence $\{a_n\}_{n=1}^{\infty}$ of *distinct* terms, and define a sequence of integers $\{n_k\}_{k=1}^{\infty}$ as follows:

$$\begin{aligned} n_1 &= \min\{n \in \mathbb{N} : a_n \in E\}, \\ n_2 &= \min\{n > n_1 : a_n \in E\}, \\ n_3 &= \min\{n > n_2 : a_n \in E\}, \\ &\vdots \\ n_k &= \min\{n > n_{k-1} : a_n \in E\}, \qquad k \geq 4, \\ &\vdots \end{aligned}$$

Since E is infinite, n_k is defined for all $k \in \mathbb{N}$. It is now clear that $E = \{a_{n_k}\}_{k=1}^{\infty}$, and so E is countable.

COROLLARY 2. *A subset of a countable set is* at most countable*, i.e. it is either countable or finite.*

The next two theorems generalize the countability of the rational numbers and the uncountability of the real numbers respectively. They are proved by the same diagonalization procedures used above, and their proofs are left to the reader.

THEOREM 2. *Let $\{E_n\}_{n=1}^{\infty}$ be a sequence of countable sets. Then $S \equiv \bigcup_{n=1}^{\infty} E_n$ is countable.*

The above theorem says that a countable union of countable sets is countable. Note that the sets E_n may overlap, but not so much as to make the union finite, since their union S contains E_1, and hence S is not finite. As an immediate corollary we may replace 'countable' with 'at most countable'.

COROLLARY 3. *An at most countable union of at most countable sets is at most countable.*

THEOREM 3. *Let A be the set of all sequences whose terms are either 0 or 1. Then A is uncountable.*

Here is one more result on countable sets that is easily proved by induction.

PROPOSITION 6. *Let A be countable and consider the n-fold product set $A^n = A \times A \times ... \times A$ defined by*

$$A^n \equiv \{(a_1, a_2, ..., a_n) : a_i \in A \text{ for } 1 \leq i \leq n\}.$$

Then A^n is countable.

Proof: Clearly $A^1 \sim A$ is countable. We now proceed by induction on n and assume that A^{n-1} is countable. Assuming that $n > 1$ we have

$$A^n = \left\{(b, c) : b \in A^{n-1} \text{ and } c \in A\right\}.$$

Now for each *fixed* $c \in A$, the set of pairs $\left\{(b, c) : b \in A^{n-1}\right\}$ is equivalent to A^{n-1} which is countable by our induction assumption. Since A is countable, we thus see that A^n is a countable union of countable sets, hence countable by Theorem 2.

CHAPTER 3

Metric spaces

There is a notion of distance between numbers in both the rational field $\mathbb{Q}$ and in the real field $\mathbb{R}$ given by the absolute value of the difference of the numbers:

$$\begin{aligned} dist\,(p,q) &= |p-q|\,, \qquad p,q \in \mathbb{Q}, \\ dist\,(x,y) &= |x-y|\,, \qquad x,y \in \mathbb{R}. \end{aligned}$$

Motivated by Pythagoras' theorem, this can be extended to complex numbers $\mathbb{C}$ by

$$\begin{aligned} dist\,(z,w) &= |z-w| = \sqrt{(x-u)^2 + (y-v)^2}, \\ \text{for } z &= x+iy \text{ and } w = u+iv \text{ in } \mathbb{C}, \end{aligned}$$

and even to points or vectors in Euclidean space:

$$\begin{aligned} dist\,(\mathbf{x},\mathbf{y}) &\equiv \|\mathbf{x}-\mathbf{y}\| = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}, \\ \text{for } \mathbf{x} &= (x_1,...,x_n) \text{ and } \mathbf{y} = (y_1,...,y_n) \text{ in } \mathbb{R}^n. \end{aligned}$$

It will eventually be important to define a notion of distance between functions, for example if f and g are continuous functions on the unit interval $[0,1]$, then we will define

$$dist\,(f,g) = \sup\left\{|f(x) - g(x)| : 0 \le x \le 1\right\}.$$

Of course at this point we don't even know if this supremum is finite, i.e. if the set in braces is bounded above, or if it is, whether or not this definition satisfies properties that we would expect of a 'distance function'. Thus we begin by setting down in as abstract a setting as possible the properties we expect of a distance function.

Definition 9. *A set X together with a function $d : X \times X \to [0,\infty)$ is said to be a* metric space*, and d is called a* metric *or* distance function *on X, provided:*

(1) $d(x,x) = 0$,
(2) $d(x,y) > 0$ *if* $x \ne y$,
(3) $d(x,y) = d(y,x)$ *for all* $x,y \in X$,
(4) *(triangle inequality)* $d(x,z) \le d(x,y) + d(y,z)$ *for all* $x,y,z \in X$.

To be precise we often write a metric space as a pair (X,d). Examples of metric spaces include $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{R}^n$ with the distance functions given above. The triangle inequality holds in $\mathbb{C}$ by Proposition 5 (1) (f). To prove that the triangle inequality holds in $\mathbb{R}^n$ we can use the Cauchy-Schwarz inequality just as we did in the proof of Proposition 5 (1) (d):

$$\begin{aligned} dist\,(x,z)^2 &= \|x-z\|^2 = \sum_{k=1}^n (x_k - z_k)^2 = \sum_{k=1}^n (x_k - y_k + y_k - z_k)^2 \\ &= \sum_{k=1}^n (x_k - y_k)^2 + 2\sum_{k=1}^n (x_k - y_k)(y_k - z_k) + \sum_{k=1}^n (y_k - z_k)^2 \\ &\le \|x-y\|^2 + 2\|x-y\|\,\|y-z\| + \|y-z\|^2 \\ &= (\|x-y\| + \|y-z\|)^2 = (dist\,(x,y) + dist\,(y,z))^2 . \end{aligned}$$

Taking square roots we obtain

$$dist\,(x,z) = \|x-z\| \le \|x-y\| + \|y-z\| = dist\,(x,y) + dist\,(y,z) . \tag{0.3}$$

We can also consider different metrics on $\mathbb{R}^n$ such as *taxicab distance*:

$$d_{taxi}\,(x,y) = \max\{|x_k - y_k| : 1 \le k \le n\} .$$

This is the shortest distance a taxi must travel to get from x to y if the taxi is restricted to proceed only vertically or horizontally, as is the case in most cities built around a rectangular grid of streets. It is not too hard an exercise to prove that $(\mathbb{R}^n, d_{taxi})$ is a metric space, i.e. that d_{taxi} satisfies the axioms in Definition 9 on the set $\mathbb{R}^n$.

An important method of constructing new metric spaces from known metric spaces is to consider *subsets*. Indeed, if (X,d) is a metric space and Y is any subset of X, then (Y,d) is *also* a metric space, as is immediately verified by restricting the points x, y, z in Definition 9 to lie in the subset Y. For example the open unit disk

$$\begin{aligned} \mathbb{D} &= \{z \in \mathbb{C} : dist\,(0,z) < 1\} \\ &= \left\{(x,y) \in \mathbb{R}^2 : \sqrt{x^2+y^2} < 1\right\} \end{aligned}$$

is a metric space with the metric $d\,(z,w) = |z-w|$. Note that the open unit disk in the complex plane $\mathbb{C}$ coincides with the open unit disk in the Euclidean plane $\mathbb{R}^2$.

The concept of a *ball* in a metric space is central to the further development of the theory of metric spaces.

Definition 10. *Let (X,d) be a metric space and suppose $x \in X$ and $r > 0$. The* ball *$B\,(x,r)$ with center x and radius r is defined to be the set of all points $y \in X$ at a distance less than r from x:*

$$B\,(x,r) = \{y \in X : d\,(x,y) < r\} .$$

One can easily verify that the collection of balls $\{B\,(x,r)\}_{x\in X, r>0}$ in a metric space (X,d) satisfies the following six properties for all $x, y \in X$:

1. $\bigcap_{r>0} B\,(x,r) = \{x\}$,
2. $\bigcup_{r>0} B\,(x,r) = X$,
3. If $0 < r \le s$, then $B\,(x,r) \subset B\,(x,s)$
4. If $y \in B\,(x,r)$, then $x \in B\,(y,r)$,
5. The set $\{r > 0 : y \in B\,(x,r)\}$ has no least element,
6. If $B\,(x,r) \cap B\,(y,s) \ne \emptyset$, then $y \in B\,(x, r+s)$.

While we will not need to know this, the six properties above characterize a metric space in the following sense. Suppose that $\{B(x,r)\}_{x\in X, r>0}$ is a collection of subsets of a set X that satisfy the six properties listed above. Define

$$d(x,y) = \inf\{r > 0 : y \in B(x,r)\}, \qquad \text{for all } x, y \in X.$$

Then it is not too hard to show that d maps $X \times X$ into $[0,\infty)$ and satisfies the four properties in Definition 9, i.e. d defines a metric or distance function on X. Moreover, one can prove that $B(x,r) = \{y \in X : d(x,y) < r\}$ for all $x \in X$ and $r > 0$, so that the initial collection of subsets $\{B(x,r)\}_{x\in X, r>0}$ are precisely the collection of balls corresponding to the metric d.

1. Topology of metric spaces

The notion of an open set is at the center of the subject of topology.

DEFINITION 11. *Let (X,d) be a metric space and suppose G is a subset of X. Then G is* open *if for every point x in G there is a* positive *radius r such that the ball $B(x,r)$ is contained in G:*

$$B(x,r) \subset G.$$

We see that the empty set $\emptyset$ is open by default (there is nothing to check). The set X is open since

$$B(x,1) \subset X, \qquad \text{for all } x \in X.$$

Any positive number would do in place of 1 as the radius above. One suspects that balls themselves are open sets, but this needs a proof which relies heavily on the triangle inequality.

LEMMA 2. *Let B be a ball in a metric space (X,d). Then B is open.*

Proof: Suppose that $B = B(y,s)$ and that $x \in B$. Then by Definition 10 we have $d(y,x) < s$. Set

$$r = s - d(x,y) > 0.$$

We claim that the ball $B(x,r)$ with center x and radius r is contained in $B(y,s)$. Draw a picture before proceeding! Indeed, if $z \in B(x,r)$ then by Definition 10 we have $d(x,z) < r$. Now we use the fact that the metric d satisfies the *triangle inequality* in Definition 9 to compute that

$$d(y,z) \le d(y,x) + d(x,z) < d(x,y) + r = s.$$

This shows that $z \in B(y,s)$ for every $z \in B(x,r)$, i.e.

$$B(x,r) \subset B(y,s).$$

Thus we have verified the condition that for every point x in $B(y,s)$ there is a *positive* radius $r = r_x$ (depending on the point x we chose in $B(y,s)$) such that the ball $B(x,r_x)$ is contained in $B(y,s)$. This proves that $B(y,s)$ is an open set.

EXERCISE 2. *Consider the Euclidean space $\mathbb{R}^2$.*

(1) *Show that the inside of the ellipse,*

$$G = \left\{(x,y) \in \mathbb{R}^2 : 4x^2 + y^2 < 1\right\},$$

is open. Hint*: If $P \equiv (x,y) \in G$, then the ball $B(P,r)$ is contained in G if*

$$r = \frac{1}{2}\left(1 - \sqrt{4x^2 + y^2}\right).$$

Indeed, if $Q = (u, v) \in B(P, r)$, then (0.3) yields

$$\begin{aligned}\sqrt{(2u)^2 + v^2} &\leq \sqrt{(2u - 2x)^2 + (v - y)^2} + \sqrt{(2x)^2 + y^2} \\ &\leq 2\sqrt{(u - x)^2 + (v - y)^2} + \sqrt{(2x)^2 + y^2} \\ &< 2r + \sqrt{(2x)^2 + y^2} = 1.\end{aligned}$$

(2) *On the other hand, show that the corresponding set*

$$F = \left\{(x, y) \in \mathbb{R}^2 : 4x^2 + y^2 \leq 1\right\},$$

defined with $\leq$ in place of $<$, is not *an open set.* Hint*: The point $P \equiv (0, 1) \in F$ but for every $r > 0$ the ball $B(P, r)$ contains the point $\left(0, 1 + \frac{r}{2}\right)$ which is not in F.*

We declare a subset F of a metric space X to be *closed* if the complement $F^c \equiv X \setminus F$ of F is an open set. For example, the set F in Exercise 2 (2) is closed, but the set G in Exercise 2 (1) is not closed.

Caution: A set may be neither open nor closed, such as the subset $[0, 1)$ of $\mathbb{R}$. Moreover, a set may be simultaneously open and closed, such as both the empty set $\emptyset$ and the entire set X in any metric space X.

Proposition 7. *Let X be a metric space.*

(1) *If $\{G_\alpha\}_{\alpha \in A}$ is a collection of open subsets, then $\bigcup_{\alpha \in A} G_\alpha$ is open,*
(2) *If $\{F_\alpha\}_{\alpha \in A}$ is a collection of closed subsets, then $\bigcap_{\alpha \in A} F_\alpha$ is closed,*
(3) *If $\{G_k\}_{k=1}^n$ is a finite collection of open subsets, then $\bigcap_{k=1}^n G_k$ is open,*
(4) *If $\{F_k\}_{k=1}^n$ is a finite collection of closed subsets, then $\bigcup_{k=1}^n F_k$ is closed.*

Proof: Suppose that G_α is open for each α and let $x \in \bigcup_{\alpha \in A} G_\alpha$. Then $x \in G_\beta$ for some β and since G_β is open, there is a ball $B(x, r) \subset G_\beta \subset \bigcup_{\alpha \in A} G_\alpha$, which shows that $\bigcup_{\alpha \in A} G_\alpha$ is open. Next suppose that F_α is closed for each α and note that if $G_\alpha = (F_\alpha)^c$, then G_α is open for each α and so $\bigcup_{\alpha \in A} G_\alpha$ is open by part (1). Thus from de Morgan's laws we have that

$$\left(\bigcap_{\alpha \in A} F_\alpha\right)^c = \bigcup_{\alpha \in A} (F_\alpha)^c = \bigcup_{\alpha \in A} G_\alpha$$

is open, so $\bigcap_{\alpha \in A} F_\alpha$ is closed by definition.

Now suppose that G_k is open for $1 \leq k \leq n$ and that $x \in \bigcup_{k=1}^n G_k$. Then there is $r_k > 0$ such that $B(x, r_k) \subset G_k$ for $1 \leq k \leq n$. It follows that if we set

$$r = \min\{r_k\}_{k=1}^n,$$

then $r > 0$ (this is where we use that the collection $\{G_k\}_{k=1}^n$ is finite) and

$$B(x, r) \subset B(x, r_k) \subset G_k, \qquad 1 \leq k \leq n.$$

Thus $B(x, r) \subset \bigcap_{k=1}^n G_k$ and this shows that $\bigcap_{k=1}^n G_k$ is open. Finally, if F_k is closed for $1 \leq k \leq n$, then $G_k = (F_k)^c$ is open and so

$$\left(\bigcup_{k=1}^n F_k\right)^c = \bigcap_{k=1}^n (F_k)^c = \bigcap_{k=1}^n G_k$$

is open by part (3). Thus $\bigcup_{k=1}^n F_k$ is closed by definition.

1.1. Subspaces. Recall that if Y is a subset of a metric space X, then we may view Y as a metric space in its own right, with metric given by that of X restricted to $Y \times Y$. The metric space (Y, d) is then called a *subspace* of (X, d), even though there is no linear structure on X. Note that if $y \in Y$ and $r > 0$, then the ball $B_Y(y, r)$ in the metric space Y satisfies

$$B_Y(y, r) = \{z \in Y : d(y, z) < r\} = B_X(y, r) \cap Y, \tag{1.1}$$

where $B_X(y, r)$ is the ball centered at y with radius r in the metric space X. Thus if E is a subset of Y, it can be considered as a subset of *either* the metric space Y *or* the metric space X. Clearly the notions of E being open or closed depend on which space is considered the ambient space. For example, if

$$E = \left\{(x, y) \in \mathbb{R}^2 : dist\left(\left(0, \frac{1}{2}\right), (x, y)\right) \leq \frac{1}{2}\right\} \setminus \{(1, 0)\}$$

is the ball center $\left(0, \frac{1}{2}\right)$ with radius $\frac{1}{2}$ together with its "boundary" except for the point $(1, 0)$, then we have

$$E \subset \mathbb{D} \subset \mathbb{R}^2.$$

Now one can show that E is a closed subset relative to the metric space $\mathbb{D}$, but it is neither open nor closed as a subset relative to the metric space $\mathbb{R}^2$. Exercise: prove this!

On the other hand, (1.1) provides the following simple connection between the open subsets relative to X and the open subsets relative Y.

THEOREM 4. *Let Y be a subset of a metric space X. Then a subset E of Y is open relative to Y* if and only if *there exists a set G open relative to X such that*

$$E = G \cap Y.$$

Proof: Suppose that E is open relative to Y. Then for each $p \in E$ there is a positive radius r_p such that $B_Y(p, r_p) \subset E$. Now set

$$G = \bigcup_{p \in E} B_X(y, r_p),$$

where we note that we are using balls B_X relative to X. Clearly G is open relative to X by Lemma 2 and Proposition 7 (1). From (1.1) we obtain

$$G \cap Y = \bigcup_{p \in E} \{B_X(y, r_p) \cap Y\} = \bigcup_{p \in E} B_Y(p, r_p),$$

and the final set is equal to E since $p \in B_Y(p, r_p) \subset E$ for each $p \in E$.

Conversely, suppose G is open relative to X and $E = G \cap Y$. Then given $p \in E$, there is $r_p > 0$ such that $B_X(p, r_p) \subset G$. From (1.1) we thus obtain

$$B_Y(p, r_p) = B_X(y, r_p) \cap Y \subset G \cap Y = E,$$

which shows that E is open relative to Y.

1.2. Limit points. In order to define the notion of limit of a function later on, we will need the idea of a *limit point* of a set. A *deleted* ball $B'(p, r)$ in a metric space is the ball $B(p, r)$ minus its center p, i.e. $B'(p, r) = B(p, r) \setminus \{p\}$.

Definition 12. *Suppose (X, d) is a metric space and that E is a subset of X. We say that $p \in X$ is a* limit point *of E if* every *deleted ball centered at p contains a point of E:*

$$B'(p, r) \cap E \neq \emptyset \qquad \textit{for all } r > 0.$$

Note the following immediate consequence of this definition:

- if p is a limit point of E then every deleted ball $B'(p, r)$ contains *infinitely* many points of E,

and so in particular E must be infinite in order to have any limits points at all. Indeed, if $B'(p, r) \cap E = \{x_j\}_{j=1}^n$ contains only n points, let $s = \min \{d(p, x_j)\}_{j=1}^n$. Then $s > 0$ and $B(p, s)$ doesn't contain any of the points $\{x_j\}_{j=1}^n$. Thus we have the contradiction $B'(p, s) \cap E = \emptyset$.

Limit points are closely related to the notion of a closed set.

Proposition 8. *A set F is closed in a metric space* if and only if *it contains all of its limit points.*

Proof: Suppose first that x is a limit point of F. Then in particular, $B(x, r) \cap F$ is nonempty for all $r > 0$, and so *no* ball $B(x, r)$ centered at x is contained in F^c. If F is closed, then F^c is open and it then follows that $x \notin F^c$. Thus $x \in F$ and we have shown that a closed set F contains all of its limit points.

Conversely, suppose that F contains all of its limit points. Pick $x \in F^c$. Since x is *not* a limit point of F, there is a deleted ball $B'(x, r)$ that does not intersect F. But $x \notin F$ as well so that $B(x, r)$ does not intersect F. Hence $B(x, r) \subset F^c$ and this shows that F^c is open, and thus that F is closed.

Definition 13. *If E is a subset of a metric space X, we define E' (the derived set of E) to be the set of all limit points of E, and we define $\overline{E}$ (the closure of E) to be $E \cup E'$, the union of E and all of its limit points.*

As a corollary to Proposition 8 we obtain the following basic theorem for the metric space $\mathbb{R}$.

Theorem 5. *Suppose that E is a nonempty subset of the real numbers $\mathbb{R}$ that is bounded above, and let $\sup E$ be the least upper bound of E. Then $\sup E$ is in $\overline{E}$, and $\sup E \in E$ if E is closed.*

Proof: Since the real numbers $\mathbb{R}$ have the Least Upper Bound Property, $z \equiv \sup E$ exists and satisfies the property that if $y < z$, then y is *not* an upper bound of E, hence there exists $x \in E$ with $y < x \leq z$. It follows that $B(z, r) \cap E \neq \emptyset$ for all $r > 0$ upon taking $y = z - r$ in the previous argument. Thus either $z \in E \subset \overline{E}$ or if not, then

$$B'(z, r) \cap E \neq \emptyset \text{ for all } r > 0,$$

in which case z is a limit point of E, hence $z \in E' \subset \overline{E}$. Finally, Proposition 8 shows that $z \in E$ if E is closed.

One might wonder if the set $\overline{E}$ contains limit points not in $\overline{E}$, or roughly speaking, if taking limit points of limit points yields new points. The answer is no, and in fact not only is $\overline{E}$ closed, it is the *smallest* closed set containing E.

Proposition 9. *If E is a subset of a metric space X, then*

$$\overline{E}=\bigcap\{F\subset X: F \text{ is closed and } E\subset F\}, \tag{1.2}$$

and $\overline{E}$ is the smallest closed set containing E.

Proof: Denote the right hand side of (1.2) by $\mathcal{E}$. Then $\mathcal{E}$ is a closed set by Proposition 7 (2). Thus by its very definition, it is the smallest closed set containing E, i.e. every other closed set F containing E contains $\mathcal{E}$. Now $\overline{E}\subset\mathcal{E}$ since by Proposition 8, every closed set F containing E also contains all the limit points E' of E.

On the other hand, if $x\notin\overline{E}$, then there exists some $r>0$ such that

$$B(x,r)\cap E=\emptyset.$$

Now $B(x,r)^c$ is closed since $B(x,r)$ is open by Lemma 2. Moreover $B(x,r)^c$ contains E and so is a candidate for the intersection defining $\mathcal{E}$. This shows that $\mathcal{E}\subset B(x,r)^c$ and in particular that $x\notin\mathcal{E}$. This proves that $\mathcal{E}\subset\overline{E}$ and completes the proof of Proposition 9.

Lemma 3. *E' is closed.*

Proof: Suppose that $z\in(E')'$ and $r>0$. Then there is $y\in B'(z,r)\cap E'$. Let $s=\min\{d(z,y),r-d(z,y)\}$. Then $s>0$ and there is $x\in B'(y,s)\cap E$. Now $x\neq z$ since otherwise $s\leq d(z,y)=d(x,y)<s$, a contradiction. Also,

$$d(z,x)\leq d(z,y)+d(y,x)<d(z,y)+s\leq r.$$

Thus $x\in B'(z,r)\cap E$ and this shows that $z\in E'$ as required.

2. Compact sets

Now we come to the single most important property that a subset of a metric space can have, namely *compactness*. In a sense, compact subsets share the most important topological properties enjoyed by *finite* sets. It turns out that the most basic of these properties is rather abstract looking at first sight, but arises so often in applications and subsequent theory that we will use it as the definition of compactness. But first we introduce some needed terminology.

Let E be a subset of a metric space X. A collection $\mathcal{G}\equiv\{G_\alpha\}_{\alpha\in A}$ of subsets G_α of X is said to be an *open cover* of E if

$$\text{each } G_\alpha \text{ is open and } E\subset\bigcup_{\alpha\in A}G_\alpha.$$

A *finite* subcover (relative to the open cover $\mathcal{G}$ of E) is a finite collection $\{G_{\alpha_k}\}_{k=1}^n$ of the open sets G_α that still covers E:

$$E\subset\bigcup_{k=1}^n G_{\alpha_k}.$$

For example, the collection $\mathcal{G}=\left\{\left(\frac{1}{n},1+\frac{1}{n}\right)\right\}_{n=1}^\infty$ of open intervals in $\mathbb{R}$ form an open cover of the interval $E=\left(\frac{1}{8},2\right)$, and $\left\{\left(\frac{1}{n},1+\frac{1}{n}\right)\right\}_{n=1}^8$ is a finite subcover. Draw a picture! However, $\mathcal{G}$ is also an open cover of the interval $E=(0,2)$ for which there is *no* finite subcover since $\frac{1}{m}\notin\left(\frac{1}{n},1+\frac{1}{n}\right)$ for all $1\leq n\leq m$.

Definition 14. *A subset E of a metric space X is* compact *if every open cover of E has a finite subcover.*

EXAMPLE 5. *Clearly every finite set is compact. On the other hand, the interval* $(0,2)$ *is not compact since* $\mathcal{G} = \left\{\left(\frac{1}{n}, 1+\frac{1}{n}\right)\right\}_{n=1}^{\infty}$ *is an open cover of* $(0,2)$ *that does* not *have a finite subcover.*

The above example makes it clear that all we need is one 'bad' cover as witness to the failure of a set to be compact. On the other hand, in order to show that an infinite set *is* compact, we must often work much harder, namely we must show that given *any* open cover, there is *always* a finite subcover. It will obviously be of great advantage if we can find simpler criteria for a set to be compact, and this will be carried out below in various situations, see e.g. Remark 8 below. For now we will content ourselves with giving one simple example of an *infinite* compact subset of the real numbers (even of the rational numbers).

EXAMPLE 6. *The set* $K \equiv \{0\} \cup \left\{\frac{1}{k}\right\}_{k=1}^{\infty}$ *is compact in* $\mathbb{R}$ *or* $\mathbb{Q}$. *Indeed, suppose that* $\mathcal{G} \equiv \{G_\alpha\}_{\alpha \in A}$ *is an open cover of* K. *Then at least one of the open sets in* $\mathcal{G}$ *contains* 0, *say* G_{α_0}. *Since* G_{α_0} *is open, there is* $r > 0$ *such that*

$$B(0,r) \subset G_{\alpha_0}.$$

Now comes the crux of the argument: there are only finitely *many points* $\frac{1}{k}$ *that lie outside* $B(0,r)$, *i.e.* $\frac{1}{k} \notin B(0,r)$ *if and only if* $k \leq \left[\frac{1}{r}\right] \equiv n$. *Now choose* G_{α_k} *to contain* $\frac{1}{k}$ *for each* k *between* 1 *and* n *inclusive (with possible repetitions). Then the finite collection of open sets* $\{G_{\alpha_0}, G_{\alpha_1}, G_{\alpha_2}, ..., G_{\alpha_n}\}$ *(after removing repetitions) constitute a finite subcover relative to the open cover* $\mathcal{G}$ *of* K. *Thus we have shown that* every *open cover of* K *has a finite subcover.*

It is instructive to observe that $K = \overline{E}$ where $E = \left\{\frac{1}{k}\right\}_{k=1}^{\infty}$ is *not* compact (since the pairwise disjoint balls $B\left(\frac{1}{k}, \frac{1}{4k^2}\right) = \left(\frac{1}{k} - \frac{1}{4k^2}, \frac{1}{k} + \frac{1}{4k^2}\right)$ cover E one point at a time). Thus the addition of the single limit point 0 to the set E resulted in making the union compact. The argument given as proof in the above example serves to illustrate the sense in which the set K is topologically 'almost' a finite set.

As a final example to illustrate the concept of compactness, we show that any *unbounded* set in a metric space fails to be compact. We say that a subset E of a metric space X is *bounded* if there is some ball $B(x,r)$ in X that contains E. So now suppose that E is unbounded. Fix a point $x \in X$ and consider the open cover $\{B(x,n)\}_{n=1}^{\infty}$ of E (this is actually an open cover of the entire metric space X). Now if there were a finite subcover, say $\{B(x,n_k)\}_{k=1}^{N}$ where $n_1 < n_2 < ... < n_N$, then because the balls are increasing,

$$E \subset \bigcup_{k=1}^{N} B(x,n_k) = B(x,n_N),$$

which contradicts the assumption that E is unbounded. We record this fact in the following lemma.

LEMMA 4. *A compact subset of a metric space is bounded.*

REMARK 8. *We can now preview one of the major themes in our development of analysis. The Least Upper Bound Property of the real numbers will lead directly to the following beautiful characterization of compactness in the metric space* $\mathbb{R}$ *of real numbers, the* Heine-Borel theorem*: a subset* K *of* $\mathbb{R}$ *is compact if and only if* K *is closed and bounded.*

Before proceeding to develop further properties of compact subsets, and their relationship to open and closed subsets, we establish a truly surprising aspect of the definition, namely that compactness is an *intrinsic* property of a set K. By this we mean:

LEMMA 5. *If $K \subset Y \subset X$ where X is a metric space, then K is compact relative to the metric space X* if and only if *it is a compact subset relative to the subspace Y.*

In particular, we can take $Y = K$ here and obtain that

- K is a compact subset of a metric space X if and only if it is compact when considered as a metric space in its own right, i.e. if and only if every cover of K by subsets of K that are open in K has a finite subcover.

This means that it makes sense to talk of a *compact* set K without reference to a larger metric space in which it is a proper subset, compare Example 6 above. Note how this contrasts with the property of a set G being open or closed, which depends heavily on the ambient metric space, see Subsection 1.1 on subspaces above.

Proof (of Lemma 5): Suppose that K is compact relative to X. We now show K is compact relative to Y. So let $\mathcal{E} \equiv \{E_\alpha\}_{\alpha \in A}$ be an open cover of K in the metric space Y. By Theorem 4 there are open sets G_α in X so that

$$E_\alpha = G_\alpha \cap Y.$$

Then $\mathcal{G} \equiv \{G_\alpha\}_{\alpha \in A}$ is an open cover of K relative to X, and since K is compact relative to X, there is a finite subcover $\{G_{\alpha_k}\}_{k=1}^n$,

$$K \subset \bigcup_{k=1}^n G_{\alpha_k}.$$

But $K \subset Y$ so that

$$K \subset K \cap Y \subset \bigcup_{k=1}^n (G_{\alpha_k} \cap Y) = \bigcup_{k=1}^n E_{\alpha_k},$$

which shows that $\{E_{\alpha_k}\}_{k=1}^n$ is a finite subcover of the open cover $\mathcal{E} \equiv \{E_\alpha\}_{\alpha \in A}$.

Conversely, suppose that K is compact relative to Y. We now show that K is compact relative to X. So let $\mathcal{G} \equiv \{G_\alpha\}_{\alpha \in A}$ be an open cover of K relative to X. If $E_\alpha = G_\alpha \cap Y$, then $\mathcal{E} \equiv \{E_\alpha\}_{\alpha \in A}$ is an open cover of K in the metric space Y. Since K is compact relative to X, there is a finite subcover $\{E_{\alpha_k}\}_{k=1}^n$. But then

$$K \subset \bigcup_{k=1}^n E_{\alpha_k} \subset \bigcup_{k=1}^n G_{\alpha_k},$$

and so $\{G_{\alpha_k}\}_{k=1}^n$ is a finite subcover of the open cover $\mathcal{G}$.

2.1. Properties of compact sets. We now prove a number of properties that hold for general compact sets. In the next subsection we will restrict attention to compact subsets of the real numbers and Euclidean spaces.

LEMMA 6. *If K is a compact subset of a metric space X, then K is a closed subset of X.*

Proof: We show that K^c is open. So fix a point $x \in K^c$. For each point $y \in K$, consider the ball $B(y, r_y)$ with

$$r_y \equiv \frac{1}{2} d(x, y). \tag{2.1}$$

Since $\{B(y, r_y)\}_{y \in K}$ is an open cover of the compact set K, there is a *finite* subcover $\{B(y_k, r_{y_k})\}_{k=1}^n$ with of course $y_k \in K$ for $1 \le k \le n$. Now by the triangle inequality and (2.1) it follows that

$$B(x, r_{y_k}) \cap B(y_k, r_{y_k}) = \emptyset, \qquad 1 \le k \le n. \tag{2.2}$$

Indeed, if the intersection on the left side of (2.2) contained a point z then we would have the contradiction

$$d(x, y_k) \le d(x, z) + d(z, y_k) < r_{y_k} + r_{y_k} = d(x, y_k).$$

Now we simply take $r = \min \{r_{y_k}\}_{k=1}^n > 0$ and note that $B(x, r) \subset B(x, r_{y_k})$ so that

$$\begin{aligned} B(x, r) \cap K &\subset B(x, r) \cap \left(\bigcup_{k=1}^n B(y_k, r_{y_k}) \right) \\ &= \bigcup_{k=1}^n \{B(x, r) \cap B(y_k, r_{y_k})\} \\ &\subset \bigcup_{k=1}^n \{B(x, r_{y_k}) \cap B(y_k, r_{y_k})\} = \bigcup_{k=1}^n \emptyset = \emptyset, \end{aligned}$$

by (2.2). This shows that $B(x, r) \subset K^c$ and completes the proof that K^c is open. Draw a picture of this proof!

LEMMA 7. *If $F \subset K \subset X$ where F is closed in the metric space X and K is compact, then F is compact.*

Proof: Let $\mathcal{G} = \{G_\alpha\}_{\alpha \in A}$ be an open cover (relative to X) of F. We must construct a finite subcover $\mathcal{S}$ of F. Now $\mathcal{G}^* = \{F^c\} \cup \mathcal{G}$ is an open cover of K. By compactness of K there is a finite subcover $\mathcal{S}^*$ of $\mathcal{G}^*$ that consists of sets from $\mathcal{G}$ and possibly the set F^c. However, if we drop the set F^c from the subcover $\mathcal{S}^*$ the resulting finite collection of sets $\mathcal{S}$ from $\mathcal{G}$ is still a cover of F (although not neccessarily of K), and provides the required finite subcover of F.

COROLLARY 4. *If F is closed and K is compact, then $F \cap K$ is compact.*

Proof: We have that K is closed by Lemma 6, and then $F \cap K$ is closed by Proposition 7 (2). Now $F \cap K \subset K$ and so Lemma 7 now shows that $F \cap K$ is compact.

REMARK 9. *With respect to unions, compact sets behave like finite sets, namely the union of finitely many compact sets is compact. Indeed, suppose K and L are compact subsets of a metric space, and let $\{G_\alpha\}_{\alpha \in A}$ be an open cover of $K \cup L$. Then there is a finite subcover $\{G_\alpha\}_{\alpha \in I}$ of K and also a (usually different) finite subcover $\{G_\alpha\}_{\alpha \in J}$ of L (here I and J are finite subsets of A). But then the union of these covers $\{G_\alpha\}_{\alpha \in I \cup J} = \{G_\alpha\}_{\alpha \in I} \cup \{G_\alpha\}_{\alpha \in J}$ is a finite subcover of $K \cup L$, which shows that $K \cup L$ is compact.*

Now we come to one of the most useful consequences of compactness in applications. A family of sets $\{E_\alpha\}_{\alpha \in A}$ is said to have the *finite intersection property* if

$$\bigcap_{\alpha \in F} E_\alpha \neq \emptyset$$

for every finite subset F of the index set A. For example the family of open intervals $\left\{\left(0, \frac{1}{n}\right)\right\}_{n=1}^\infty$ has the finite intersection property despite the fact that the sets have no element in common: $\bigcap_{n=1}^{\infty} \left(0, \frac{1}{n}\right) = \emptyset$. The useful consequence of compactness referrred to above is that this *cannot* happen for compact subsets!

THEOREM 6. *Suppose that $\{K_\alpha\}_{\alpha \in A}$ is a family of* compact *sets with the finite intersection property. Then*

$$\bigcap_{\alpha \in A} K_\alpha \neq \emptyset.$$

Proof: Fix a member K_{α_0} of the family $\{K_\alpha\}_{\alpha \in A}$. Assume in order to derive a contradiction that *no* point of K_{α_0} belongs to every K_α. Then the open sets $\{K_\alpha^c\}_{\alpha \in A \setminus \{\alpha_0\}}$ form an open cover of K_{α_0}. By compactness, there is a finite subcover $\{K_\alpha^c\}_{\alpha \in F \setminus \{\alpha_0\}}$ with F finite, so that

$$K_{\alpha_0} \subset \bigcup_{\alpha \in F \setminus \{\alpha_0\}} K_\alpha^c,$$

i.e.

$$K_{\alpha_0} \cap \bigcap_{\alpha \in F \setminus \{\alpha_0\}} K_\alpha = \emptyset,$$

which contradicts our assumption that the finite intersection property holds.

COROLLARY 5. *If $\{K_n\}_{n=1}^\infty$ is a nonincreasing sequence of nonempty compact sets. i.e. $K_{n+1} \subset K_n$ for all $n \geq 1$, then*

$$\bigcap_{n=1}^{\infty} K_n \neq \emptyset.$$

THEOREM 7. *If E is an infinite subset of a compact set K, then E has a limit point in K.*

Proof: Suppose, in order to derive a contradiction, that *no* point of K is a limit point of E. Then for each $z \in K$, there is a ball $B(z, r_z)$ that contains *at most* one point of E (namely z if z is in E). Thus it is not possible for a finite number of these balls $B(z, r_z)$ to cover the infinite set E. Thus $\{B(z, r_z)\}_{z \in K}$ is an open cover of K that has no finite subcover (since a finite subcover cannot cover even the subset E of K). This contradicts the assumption that K is compact.

There is a converse to this theorem that leads to the following characterization of compactness in a general metric space.

THEOREM 8. *A metric space (X, d) is compact if and only if every infinite subset of X has a limit point in X.*

Proof: The 'only if' statement is Theorem 7. The proof of the 'if' statement is a bit delicate, and we content ourselves with a mere sketch here. First we note that X has a countable *dense subset* E, i.e. every open subset G contains a point of E. Indeed, for each $n \in \mathbb{N}$ there exists a finite set of balls $\left\{B_K\left(x_k^n, \frac{1}{n}\right)\right\}_{k=1}^{N_n}$ that cover X. To see this we inductively define x_k^n so that $d\left(x_i^n, x_k^n\right) \geq \frac{1}{n}$ for all $1 \leq i < k$, and note that the process must terminate since otherwise $\{x_i^n\}_{i=1}^{\infty}$ would be an infinite subset of X with no limit point, a contradiction. The set $E = \bigcup_{n=1}^{\infty} \{x_k^n\}_{k=1}^{N_n}$ is then countable and dense in K. Second we use this to construct a countable *base* for X, i.e. a countable collection of open sets $\mathcal{B} = \{B_n\}_{n=1}^{\infty}$ such that for every open set G and $z \in G$ there is $n \geq 1$ such that $z \in B_n \subset G$. Indeed, if E is a countable dense subset, then $\mathcal{B} = \{B(x, r) : x \in E, r \in \mathbb{Q} \cap (0, 1)\}$ is a countable base.

Now suppose that $\{G_\alpha\}_{\alpha \in A}$ is an open cover of X. For each $x \in X$ there is an index $\alpha \in A$ and a ball $B_x \in \mathcal{B}$ such that

$$x \in B_x \subset G_\alpha. \tag{2.3}$$

Note that the axiom of choice is not needed here since $\mathcal{B}$ is countable, hence well-ordered. If we can show that the open cover $\widetilde{\mathcal{B}} = \{B_x : x \in X\}$ has a finite subcover, then (2.3) shows that $\{G_\alpha\}_{\alpha \in A}$ has a finite subcover as well. So it remains to show that $\widetilde{\mathcal{B}}$ has a finite subcover. Relabel the open cover $\widetilde{\mathcal{B}}$ as $\widetilde{\mathcal{B}} = \{B_n\}_{n=1}^{\infty}$. Assume, in order to derive a contradiction, that $\widetilde{\mathcal{B}}$ has no finite subcover. Then the sets

$$F_N = X \setminus \left(\bigcup_{k=1}^{N} B_n\right)$$

are *nonempty* closed sets that are decreasing, i.e. $F_{N+1} \subset F_N$, and that have *empty* intersection. Thus if we choose $x_N \in F_N$ for each N, the set $E = \bigcup_{N=1}^{\infty} \{x_N\}$ must be an infinite set, and so has a limit point $x \in X$. But then the fact that the F_N are closed and decreasing implies that $x \in F_N$ for all N, the desired contradiction.

2.2. Compact subsets of Euclidean space. The Least Upper Bound Property of the real numbers plays a crucial role in the proof that closed bounded intervals are compact.

THEOREM 9. *The closed interval* $[a, b]$ *is compact (with the usual metric) for all* $a < b$.

We give two proofs of this basic theorem. The second proof will be generalized to prove that closed bounded rectangles in $\mathbb{R}^n$ are compact.

Proof #1: Assume for convenience that the interval is the closed unit interval $[0, 1]$, and suppose that $\{G_\alpha\}_{\alpha \in A}$ is an open cover of $[0, 1]$. Now $1 \in G_\beta$ for some $\beta \in A$ and thus there is $r > 0$ such that $(1 - r, 1 + r) \subset G_\beta$. With $a = 1 + \frac{r}{2} > 1$ it follows that $\{G_\alpha\}_{\alpha \in A}$ is an open cover of $[0, a]$. Now define

$$E = \{x \in [0, a] : \text{the interval } [0, x] \text{ has a finite subcover}\}.$$

We have E is nonempty ($0 \in E$) and bounded above (by a). Thus $\lambda \equiv \sup E$ exists. We claim that $\lambda > 1$. Suppose for the moment that this has been proved. Then 1 cannot be an upper bound of E and so there is some $\sigma \in E$ satisfying

$$1 < \sigma \leq \lambda.$$

Thus by the definition of the set E it follows that $[0, \sigma]$ has a finite subcover, and hence so does $[0, 1]$, which completes the proof of the theorem.

Now suppose, in order to derive a contradiction, that $\lambda \leq 1$. Then there is some open set G_γ with $\gamma \in A$ and also some $s > 0$ such that

$$(\lambda - s, \lambda + s) \subset G_\gamma.$$

Now by the definition of least upper bound, there is some $x \in E$ satisfying $\lambda - s < x \leq \lambda$, and by taking s less than $a - 1$ we can also arrange to have

$$\lambda + s \leq 1 + s < a.$$

Thus there is a finite subcover $\{G_{\alpha_k}\}_{k=1}^n$ of $[0, x]$, and if we include the set G_γ with this subcover we get a finite subcover of $\left[0, \lambda + \frac{s}{2}\right]$. This shows that $\lambda + \frac{s}{2} \in E$, which contradicts our assumption that λ is an upper bound of E, and completes the proof of the theorem.

Proof #2: Suppose, in order to derive a contradiction, that there is an open cover $\{G_\alpha\}_{\alpha \in A}$ of $[a, b]$ that has *no* finite subcover. Then at least one of the two intervals $\left[a, \frac{a+b}{2}\right]$ and $\left[\frac{a+b}{2}, b\right]$ fails to have a finite subcover. Label it $[a_1, b_1]$ so that

$$\begin{aligned} a &\leq a_1 < b_1 \leq b, \\ b_1 - a_1 &= \frac{1}{2}\delta, \end{aligned}$$

where $\delta = b - a$. Next we note that at least one of the two intervals $\left[a_1, \frac{a_1+b_1}{2}\right]$ and $\left[\frac{a_1+b_1}{2}, b_1\right]$ fails to have a finite subcover. Label it $[a_2, b_2]$ so that

$$\begin{aligned} a &\leq a_1 \leq a_2 < b_2 \leq b_1 \leq b, \\ b_2 - a_2 &= \frac{1}{4}\delta. \end{aligned}$$

Continuing in this way we obtain for each $n \geq 2$ an interval $[a_n, b_n]$ such that

$$\begin{aligned} a &\leq a_1 \leq ...a_{n-1} \leq a_n < b_n \leq b_{n-1}... \leq b_1 \leq b, \\ b_n - a_n &= \frac{1}{2^n}\delta. \end{aligned} \tag{2.4}$$

Now let $E = \{a_n : n \geq 1\}$ and set $x \equiv \sup E$. From (2.4) we obtain that each b_n is an upper bound for E, hence $x \leq b_n$ and we have

$$a \leq a_n \leq x \leq b_n \leq b, \text{ for all } n \geq 1,$$

i.e. $x \in [a_n, b_n]$ for all $n \geq 1$. Now $x \in [a, b]$ and so there is $\beta \in A$ and $r > 0$ such that

$$(x - r, x + r) \subset G_\beta.$$

By the Archimedean property of $\mathbb{R}$ we can choose $n \in \mathbb{N}$ so large that $\frac{1}{r} < n < 2^n$ (it is easy to prove $n < 2^n$ for all $n \in \mathbb{N}$ by induction), and hence

$$[a_n, b_n] \subset (x - r, x + r) \subset G_\beta.$$

But this contradicts our construction that $[a_n, b_n]$ has no finite subcover, and completes the proof of the theorem.

Corollary 6. *A subset K of the real numbers $\mathbb{R}$ is compact* if and only if *K is closed and bounded.*

Proof: Suppose that K is compact. Then K is bounded by Lemma 4 and is closed by Lemma 6. Conversely if K is bounded, then $K \subset [-a, a]$ for some $a > 0$. Now $[-a, a]$ is compact by Theorem 9, and if K is closed, then Lemma 7 shows that K is compact.

Proof #2 of Theorem 9 is easily adapted to prove that closed rectangles

$$R = \prod_{k=1}^{n} [a_k, b_k] = [a_1, b_1] \times ... \times [a_n, b_n]$$

in $\mathbb{R}^n$ are compact.

THEOREM 10. *The closed rectangle* $R = \prod_{k=1}^{n} [a_k, b_k]$ *is compact (with the usual metric) for all* $a_k < b_k$, $1 \leq k \leq n$.

Proof: Here is a brief sketch of the proof. Suppose, in order to derive a contradiction, that there is an open cover $\{G_\alpha\}_{\alpha \in A}$ of R that has *no* finite subcover. It is convenient to write R as a product of closed intervals with superscripts instead of subscripts: $R = \prod_{k=1}^{n} \left[a^k, b^k\right]$. Now divide R into 2^n congruent closed rectangles. At least one of them fails to have a finite subcover. Label it $R_1 \equiv \prod_{k=1}^{n} \left[a_1^k, b_1^k\right]$, and repeat the process to obtain a sequence of decreasing rectangles $R_m \equiv \prod_{k=1}^{n} \left[a_m^k, b_m^k\right]$ with

$$\begin{aligned} a^k &\leq a_1^k \leq ...a_{m-1}^k \leq a_m^k < b_m^k \leq b_{m-1}^k ... \leq b_1^k \leq b^k, \\ b_m^k - a_m^k &= \frac{1}{2^m} \delta^k, \end{aligned}$$

where $\delta^k = b^k - a^k$, $1 \leq k \leq n$. Then if we set $x^k = \sup \left\{a_m^k : m \geq 1\right\}$ we obtain that $x = \left(x^1, ..., x^n\right) \in R_m \subset R$ for all m. Thus there is $\beta \in A$, $r > 0$ and $m \geq 1$ such that

$$R_m \subset B(x, r) \subset G_\beta,$$

contradicting our construction that R_m has no finite subcover.

THEOREM 11. *Let K be a subset of Euclidean space $\mathbb{R}^n$. Then the following three conditions are equivalent:*

(1) *K is closed and bounded;*
(2) *K is compact;*
(3) *every infinite subset of K has a limit point in K.*

Proof: We prove that (1) implies (2) implies (3) implies (1). If K is closed and bounded, then it is contained in a closed rectangle R, and is thus compact by Theorem 10 and Lemma 7. If K is compact, then every infinite subset of K has a limit point in K by Theorem 7. Finally suppose that every infinite subset of K has a limit point in K. Of course Theorem 8 implies that K is compact, hence closed and bounded by Lemmas 6 and 4, but in Euclidean space there is a much simpler proof that avoids the use of Theorem 8.

Suppose first, in order to derive a contradiction, that K is not bounded. Then there is a sequence $\{x_k\}_{k=1}^{\infty}$ of points in K with $|x_k| \geq k$ for all k. Clearly the set of points in $\{x_k\}_{k=1}^{\infty}$ is an infinite subset E of K but has no limit point in $\mathbb{R}^n$, hence not in K either. Suppose next, in order to derive a contradiction, that K is not closed. Then there is a limit point x of K that is not in K. Thus each deleted ball $B'\left(x, \frac{1}{k}\right)$ contains some point x_k from K. Again it is clear that the set of points in the sequence $\{x_k\}_{k=1}^{\infty}$ is an infinite subset of K but contains no limit point in K since its only limit point is x and this is not in K.

COROLLARY 7. *Every bounded infinite subset of $\mathbb{R}^n$ has a limit point in $\mathbb{R}^n$.*

3. Fractal sets

We say that a subset E of a Euclidean space $\mathbb{R}^n$ is a *fractal set* if it replicates under dilation and translation/rotation in the following way: there are positive integers k and m such that

$$kE = E_1 \cup E_2 \cup ... \cup E_m, \tag{3.1}$$

where kE is a dilation of E by factor k,

$$kE \equiv \{kx : x \in E\},$$

each E_j is a translation and rotation of E by some vector a_j and rotation matrix M_j,

$$E_j \equiv \{M_j (x + a_j) : x \in E\},$$

and finally where the sets E_j are pairwise disjoint (sometimes we will relax this condition somewhat to require some notion of 'essentially' pairwise disjoint). We will refer to the number $\alpha \equiv \frac{\ln m}{\ln k}$ as the *fractal dimension* of E. This terminology is explained below.

The simplest example of a fractal is the unit half open half closed cube $\mathbb{I}_n$ in $\mathbb{R}^n$:

$$\begin{aligned}
\mathbb{I}_1 &= [0,1), \\
\mathbb{I}_2 &= [0,1) \times [0,1), \\
\mathbb{I}_n &= [0,1)^n = \prod_{j=1}^{n} [0,1).
\end{aligned}$$

With $E = \mathbb{I}_n$, $k = 2$ and $m = 2^n$ we have,

$$\begin{aligned}
kE &= 2\mathbb{I}_n = [0,2)^n \\
&= \bigcup_{(\ell_1,...,\ell_n)\in\{0,1\}^n} ([0,1)^n + (\ell_1, ..., \ell_n)) \\
&= \bigcup_{j=1}^{2^n} (\mathbb{I}_n + a_j) = \bigcup_{j=1}^{m} E_j,
\end{aligned}$$

where $\{a_j\}_{j=1}^{2^n}$ is an enumeration of the 2^n sequences $(\ell_1, ..., \ell_n)$ of $0'$s and 1's having length n. Note that if we let k denote an integer larger than 2, then we would have

$$kE = \bigcup_{j=1}^{m} E_j$$

with $m = k^n$. Thus the quantity which remains *invariant* in these calculations is the exponent n satisfying $m = k^n$ or

$$n = \log_k m = \frac{\ln m}{\ln k}.$$

Note that the compact set $\overline{\mathbb{I}_n} = [0,1]^n$ also satisfies (3.1) with the same translations, but where the E_j overlap on edges. As n is the dimension of the cube $\mathbb{I}_n$, we will more generally refer to the quantity

$$\alpha \equiv \log_k m = \frac{\ln m}{\ln k}$$

associated to a fractal set E as the *fractal dimension* of E. It can be shown that if E satisfies two different pairwise disjoint replications $k_1 E = \bigcup_{j=1}^{m_1} E_j$ and $k_2 E = \bigcup_{j=1}^{m_2} E_j$, then $\alpha = \log_{k_1} m_1 = \log_{k_2} m_2$ is independent of the replication and depends only on E.

3.1. The Cantor set. We now construct our first nontrivial fractal, the Cantor middle thirds set (1883). It turns out to have fractional dimension. We start with the closed unit interval $I = I^0 = [0,1]$. Now remove the open middle third $\left(\frac{1}{3}, \frac{2}{3}\right)$ of length $\frac{1}{3}$ and denote the two remaining closed intervals of length $\frac{1}{3}$ by $I_1^1 = \left[0, \frac{1}{3}\right]$ and $I_2^1 = \left[\frac{2}{3}, 1\right]$. Then remove the open middle third $\left(\frac{1}{9}, \frac{2}{9}\right)$ of length $\frac{1}{3^2}$ from $I_1^1 = \left[0, \frac{1}{3}\right]$ and denote the two remaining closed intervals of length $\frac{1}{3^2}$ by I_1^2 and I_2^2. Do the same for I_2^1 and denote the two remaining closed intervals by I_3^2 and I_4^2.

Continuing in this way, we obtain at the k^{th} generation, a collection $\left\{I_j^k\right\}_{j=1}^{2^k}$ of 2^k pairwise disjoint closed intervals of length $\frac{1}{3^k}$. Let $K_k = \bigcup_{j=1}^{2^k} I_j^k$ and set

$$E = \bigcap_{k=1}^{\infty} K_k = \bigcap_{k=1}^{\infty} \left(\bigcup_{j=1}^{2^k} I_j^k \right).$$

Now by Proposition 7 each set K_k is closed, and hence so is the intersection E. Then E is *compact* by Corollary 6. It also follows from Corollary 5 that E is *nonempty*. Next we observe that by its very construction, E is a fractal satisfying the replication identity

$$3E = E \cup (E+2) = E_1 \cup E_2.$$

Thus the fractal dimension α of the Cantor set E is $\frac{\ln 2}{\ln 3}$. Moreover, E has the property of being *perfect*.

Definition 15. *A subset E of a metric space X is* perfect *if E is closed and every point in E is a limit point of E.*

To see that the Cantor set is perfect, pick $x \in E$. For each $k \geq 1$ the point x lies in exactly one of the closed intervals I_j^k for some j between 1 and 2^k. Since the length of I_j^k is positive, in fact $\frac{1}{3^k} > 0$, it is possible to choose a point $x_k \in I_j^k \setminus \{x\}$. Now the set of points in the sequence $\{x_k\}_{k=1}^{\infty}$ is an infinite subset of E and clearly has x as a limit point. This completes the proof that the Cantor set E is perfect.

By summing the lengths of the removed open middle thirds, we obtain

$$\text{`length'}\left([0,1] \setminus E\right) = \frac{1}{3} + \frac{2}{3^2} + \frac{2^2}{3^3} + ... = 1,$$

and it follows that E is nonempty, compact and has 'length' $1 - 1 = 0$. Another way to exhibit the same phenomenon is to note that for each $k \geq 1$ the Cantor set E is a subset of the closed set K_k which is a union of 2^k intervals each having length $\frac{1}{3^k}$. Thus the 'length' of K_k is $2^k \frac{1}{3^k} = \left(\frac{2}{3}\right)^k$, and the 'length' of E is at most

$$\inf \left\{ \left(\frac{2}{3}\right)^k : k \geq 1 \right\} = 0.$$

In contrast to this phenomenon that the 'length' of E is quite small, the cardinality of E is quite large, namely E is uncountable, as is every nonempty perfect subset of a metric space with the Heine-Borel property: every closed and bounded subset is compact. We will need the following easily proved fact:

- In any metric space X, the closure $\overline{B(x,r)}$ of the ball $B(x,r)$ satisfies
$$\overline{B(x,r)} \subset \{y \in X : d(x,y) \leq r\}.$$

THEOREM 12. *Suppose X is a metric space in which every closed and bounded subset is compact. Then every nonempty perfect subset of X is uncountable.*

Proof: Suppose that P is a nonempty perfect subset of X. Since P has a limit point it must be infinite. Now assume, in order to derive a contradiction, that P is countable, say $P = \{x_n\}_{n=1}^\infty$. Start with any point $y_1 \in P$ that is not x_1 and the ball $B_1 \equiv B(y_1, r_1)$ where $r_1 = \frac{d(x_1,y_1)}{2}$. We have

$$B_1 \cap P \neq \emptyset \text{ and } x_1 \notin \overline{B_1}.$$

Then there is a point $y_2 \in B_1' \cap P$ that is not x_2 and so we can choose a ball B_2 such that

$$B_2 \cap P \neq \emptyset \text{ and } x_2 \notin \overline{B_2} \text{ and } \overline{B_2} \subset B_1.$$

Indeed, we can take $B_2 = B(y_2, r_2)$ where $r_2 = \frac{\min\{d(x_2,y_2), r_1 - d(y_1,y_2)\}}{2}$. Continuing in this way we obtain balls B_k satisfying

$$B_k \cap P \neq \emptyset \text{ and } x_k \notin \overline{B_k} \text{ and } \overline{B_k} \subset B_{k-1}, \qquad k \geq 2.$$

Now we use the hypothesis that every closed and bounded set in X is compact. It follows that each closed set $\overline{B_k} \cap P$ is nonempty and compact, and so by Corollary 5 we have

$$\bigcap_{k=1}^\infty \left(\overline{B_k} \cap P\right) \neq \emptyset, \qquad \text{say } x \in \left(\bigcap_{k=1}^\infty \overline{B_k}\right) \cap P.$$

However, by construction we have $x_n \notin \overline{B_n}$ for all n and since the sets $\overline{B_n}$ are decreasing, we see that $x_n \notin \bigcap_{k=1}^\infty \overline{B_k}$ for all n; hence $x \neq x_n$ for all n. This contradicts $P = \{x_n\}_{n=1}^\infty$ and completes the proof of the theorem.

3.2. The Sierpinski triangle, Cantor dust and von Koch snowflake. The *Sierpinski triangle* is a plane version of the Cantor set. Begin with the unit solid equilateral triangle $T = T^0 = \triangle\left((0,0), (1,0), \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)\right)$ whose edges of length 1 join the three points $(0,0), (1,0), \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ in the plane. Divide T^0 into four congruent equilateral triangles with edgelength $\frac{1}{2}$ by joining the midpoints of the three edges of T^0. Remover the center (upside down) open equilateral triangle to leave three closed equilateral triangles T_1^1, T_2^1, T_3^1 of edgelength $\frac{1}{2}$. Repeat this construction to obtain at the k^{th} generation, a collection $\left\{T_j^k\right\}_{j=1}^{3^k}$ of 3^k pairwise disjoint closed solid equilateral triangles of edgelength $\frac{1}{2^k}$. Let $K_k = \bigcup_{j=1}^{3^k} T_j^k$ and set

$$S = \bigcap_{k=1}^{\infty} K_k = \bigcap_{k=1}^{\infty} \left(\bigcup_{j=1}^{3^k} T_j^k \right).$$

Then the Sierpinski triangle S is a nonempty compact perfect subset of $\mathbb{R}^2$ that has 'area' equal to 0. Moreover S is a fractal satisfying the replication identity

$$2S = S \cup (S + (1,0)) \cup \left(S + \left(1, \sqrt{3}\right)\right) = S_1 \cup S_2 \cup S_3,$$

and so has fractal dimension $\frac{\ln 3}{\ln 2}$.

The *Cantor dust* is another plane version of the Cantor set, this time with fractal dimension 1. From the unit closed square $[0,1]^2$ remove everything but the four closed squares of side lenth $\frac{1}{4}$ at the corners of $[0,1]^2$, i. e. the squares $\left[0, \frac{1}{4}\right]^2$, $\left[\frac{3}{4}, 1\right] \times \left[0, \frac{1}{4}\right]$, $\left[\frac{3}{4}, 1\right]^2$ and $\left[0, \frac{1}{4}\right] \times \left[\frac{3}{4}, 1\right]$. Then repeat this procedure with these four smaller squares and continue ad infinitum. The 'dust' D that remains is a nonempty perfect compact subset of the plane satisfying the replication formula

$$D = \frac{1}{4} D \cup \frac{1}{4} (D + (3,0)) \cup \frac{1}{4} (D + (0,3)) \cup \frac{1}{4} (D + (3,3)).$$

Thus D has fractal dimension $\frac{\ln 4}{\ln 4} = 1$. The set D is in stark contrast to the segment $\left\{(x,0) \in \mathbb{R}^2 : 0 \leq x \leq 1\right\}$ in the plane that also has fractal dimension 1.

Finally, the *von Koch snowflake* (1904) is a bit harder to construct rigorously at this stage, although we will return to it later on after we have studied the concept of uniform convergence. For now we simply describe the snowflake-shaped curve informally. Begin with the line segment L^0 joining the points $(0,0)$ and $(1,0)$ along the x -axis. It is a segment of length 1 that looks like ___. Now divide the segment L^0 into three congruent closed line segments of length $\frac{1}{3}$ that each look like _, and denote the first and last of these by L_1^1 and L_4^1 respectively. Now replace the middle segment with the two segments L_2^1 joining $\left(\frac{1}{3}, 0\right)$ to $\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$ and L_3^1 joining $\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$ to $\left(\frac{2}{3}, 0\right)$. Thus the middle third segment _ has been replaced with a 'hat' that looks like $\wedge$, which together with the removed middle third makes an equilateral triangle of side length $\frac{1}{3}$. The four segments $\left\{L_j^1\right\}_{j=1}^4$ form a connected polygonal path that looks like _ $\wedge$ _ where each of the four segments has length $\frac{1}{3}$. Now we continue by replacing each of the four segments L_j^1 of length $\frac{1}{3}$ by the polygonal path of four segments of length $\frac{1}{3^2}$ obtained by removing the middle third of L_j^1 and replacing it by two equal length segments as above. Repeat this construction to obtain at the k^{th} generation, a polygonal path consisting of 4^k closed segments$\left\{L_j^k\right\}_{j=1}^{4^k}$ of length $\frac{1}{3^k}$ each. Denote this polygonal 'snowflake-shaped' path by P_k.

We now define the von Koch snowflake K to be the 'limit' of the polygonal paths P_k as $k \to \infty$. A more precise definition is this:

- K consists of all $(x,y) \in \mathbb{R}^2$ such that for every $\varepsilon > 0$ there is N satisfying

$$B\left((x,y), \varepsilon\right) \cap P_k \neq \emptyset, \qquad \text{for all } k \geq N.$$

In other words, K is the set of points in the plane such that every ball centered at the point intersects *all* of the polygonal paths from some index on. One can show that K is a compact subset of the plane that satisfies the replication identity

$$3K = K_1 \cup K_2 \cup K_3 \cup K_4,$$

where each K_j is a translation and rotation of K; moreover two different K_j intersect in at most one point. It follows that K has fractal dimension $\frac{\ln 4}{\ln 3}$. Later we will show that K is the image of a continuous curve with no tangent at any point, and infinite length between any two distinct points on it.

Here is a table of some of the fractals we constructed above. The matrices M_2 and M_3 are plane rotations through angles of $\frac{\pi}{4}$ and $-\frac{\pi}{4}$ respectively.

Fractal Set F	**Replication formula**	**Dimension**	
E	$F = \frac{1}{3}F \cup \frac{1}{3}(F+2)$	$\frac{\ln 2}{\ln 3}$	0.63093
$[0,1]$	$F = \frac{1}{2}F \cup \frac{1}{2}(F+1)$	1	1
D	$F = \frac{1}{4}F \cup \frac{1}{4}(F+(3,0)) \cup \frac{1}{4}(F+(0,3)) \cup \frac{1}{4}(F+(3,3))$	1	1
K	$F = \frac{1}{3}F \cup \frac{1}{3}(M_2F+(1,0)) \cup \frac{1}{3}\left(M_3F+\left(\frac{3}{2},\frac{\sqrt{3}}{2}\right)\right) \cup \frac{1}{3}(F+(2,0))$	$\frac{\ln 4}{\ln 3}$	1.2619
S	$F = \frac{1}{2}F \cup \frac{1}{2}(F+(1,0)) \cup \frac{1}{2}\left(F+\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right)\right)$	$\frac{\ln 3}{\ln 2}$	1.5850
$[0,1]^2$	$F = \frac{1}{2}F \cup \frac{1}{2}(F+(1,0)) \cup \frac{1}{2}(F+(0,1)) \cup \frac{1}{2}(F+(1,1))$	2	2

3.3. Similarities: A fixed point theorem. Each of the fractals F considered in the previous subsection satisfies a replication formula of the form

$$F = S_1(F) \cup S_2(F) \cup \ldots \cup S_m(F), \tag{3.2}$$

where $m \geq 2$ and each S_j is a similarity transformation in $\mathbb{R}^n$, i.e. a composition of a translation, rotation and a dilation with ratio $0 < r_j < 1$. Moreover, in all of our examples each S_j is a dilation with the same ratio $0 < r < 1$. Our next theorem shows that no matter what similarities we consider with positive dilation ratios less than 1, there is always a nonempty compact set F that satisfies (3.2), and furthermore F is uniquely determined by (3.2). Note that we are not requiring that the sets S_jF be pairwise disjoint here. We call a nonempty set F satisfying (3.2) a *self-similar* set. If all the dilations have the same ratio, we say that F is a *fractal* set. The sets listed in the table above are all compact fractal sets.

In order to prove uniqueness in our theorem on self-similarity we will use a special metric space whose elements are the nonempty compact subsets of $\mathbb{R}^n$. For $n \in \mathbb{N}$ let

$$\mathcal{X}^n \equiv \{K \subset \mathbb{R}^n : K \text{ is nonempty and compact}\}.$$

Given a pair of compact sets K, L in $\mathcal{X}^n$ we define a distance between them by

$$d(K,L) \equiv \inf\{\delta > 0 : K \subset L_\delta \text{ and } L \subset K_\delta\}, \tag{3.3}$$

where $K_\delta \equiv \{x \in \mathbb{R}^n : dist(x, K) < \delta\}$ and $dist(x, K) \equiv \inf_{y \in K} |x - y|$ is the usual distance between a point x and a set K. It is a straightforward exercise to prove that $d : \mathcal{X}^n \times \mathcal{X}^n \to [0, \infty)$ satisfies the properties of a metric as in Definition 9.

Exercise 3. *Prove that d is a metric on $\mathcal{X}^n$. Why can't we allow $\emptyset \in \mathcal{X}^n$?* Hint*: To see that $d(K, L) > 0$ if $K \neq L$, we may suppose that $x \in K \setminus L$. Then the open cover $\left\{B\left(y, \frac{d(x,y)}{2}\right)\right\}_{y \in L}$ of the compact set L has a finite subcover $\left\{B\left(y_j, \frac{d(x,y_j)}{2}\right)\right\}_{j=1}^{N}$. If $r = \min_{1 \le j \le N} \frac{d(x,y_j)}{2}$, then $r > 0$ and $B(x, r) \cap L = \emptyset$. It follows that $d(K, L) \ge d(x, L) \ge r > 0$. To see why we can't allow $\emptyset \in \mathcal{X}^n$, show that $d(\emptyset, \{x\}) = \infty$ for any $x \in X$.*

The space $\mathcal{X}^n$ can also be viewed as an extension of $\mathbb{R}^n$ via the map that takes x in $\mathbb{R}^n$ to the compact set $\{x\}$ in $\mathcal{X}^n$. This map is actually an *isometry*, meaning that it preserves distances:

$$dist_{\mathbb{R}^n}(x, y) \equiv |x - y| = d(\{x\}, \{y\}).$$

We will construct a solution to (3.2) using the finite intersection property of compact sets, and then prove uniqueness using a fixed point argument in the metric space $(\mathcal{X}^n, d)$. To see the connection with a fixed point, define for any set F,

$$\widetilde{S}(F) \equiv \bigcup_{j=1}^{m} S_j(F) \tag{3.4}$$

to be the right hand side of (3.2). Note that S_j takes balls to balls, hence bounded sets to bounded sets and open sets to open sets, hence also closed sets to closed sets. By Theorem 11 it follows that S_j takes compact sets to compact sets, and hence so does $\widetilde{S}$. Thus $\widetilde{S}$ maps the metric space $\mathcal{X}^n$ into itself, and moreover, a set $F \in \mathcal{X}^n$ is self-similar if and only if F is a *fixed point* of $\widetilde{S}$, i.e. $\widetilde{S}(F) = F$.

Here is the theorem on existence of self-similar sets, which exhibits a simple classification, in terms of similarity transformations, of these very complex looking sets. It was B. Mandelbrot (1977) who brought the world's attention to the fact that much of the seeming complexity in nature is closely related to self-similarity - plants, trees, shells, rivers, coastlines, mountain ranges, clouds, lightning, etc.

Theorem 13. *For $1 \le j \le m$ suppose that S_j is a similarity transformation on $\mathbb{R}^n$ with dilation ratio $0 < r_j < 1$. Then there is a unique nonempty compact subset F of $\mathbb{R}^n$ satisfying (3.2).*

Proof: We begin by choosing a *closed* ball $B = B_R = \{x \in \mathbb{R}^n : |x| \le R\}$ so large that

$$S_j(B) \subset B, \qquad 1 \le j \le m. \tag{3.5}$$

Since

$$|S_j(x)| \le |S_j(x) - S_j(0)| + |S_j(0)| \le r_j |x| + |S_j(0)|,$$

it suffices to take

$$R = \frac{\max_{1 \le i \le m} |S_i(0)|}{1 - \max_{1 \le i \le m} r_i},$$

so that if $x \in B_R$, then

$$\begin{aligned} |S_j(x)| &\leq \left(\max_{1\leq i\leq m} r_i\right) R + \max_{1\leq i\leq m} |S_i(0)| \\ &= \left(\max_{1\leq i\leq m} r_i\right) R + \left(1 - \max_{1\leq i\leq m} r_i\right) R = R. \end{aligned}$$

A trivial property of the set mapping $\widetilde{S}$ is monotonicity:

$$\widetilde{S}(E) \subset \widetilde{S}(F) \text{ if } E \subset F. \tag{3.6}$$

A less obvious property, which will be used to prove the uniqueness assertion in Theorem 13, is a contractive inequality relative to the distance d introduced above for the metric space $\mathcal{X}^n$:

$$d\left(\widetilde{S}(A), \widetilde{S}(B)\right) \leq r d(A, B), \qquad A, B \in \mathcal{X}^n, \tag{3.7}$$

where $r \equiv \max_{1\leq i\leq m} r_i$. To see (3.7), it suffices by symmetry to show that

$$\widetilde{S}(A) \subset \left(\widetilde{S}(B)\right)_{r(d(A,B)+\varepsilon)}, \qquad \text{for all } \varepsilon > 0. \tag{3.8}$$

So pick $\zeta \in \widetilde{S}(A)$, i.e. $\zeta = S_j(x) \in S_j(A)$ for some $x \in A$ and $1 \leq j \leq m$. Now for any $\varepsilon > 0$ we know that $x \in A \subset B_{d(A,B)+\varepsilon}$ so that there is $y \in B$ satisfying $|x - y| < d(A, B) + \varepsilon$. Then

$$\eta = S_j(y) \in S_j(B) \subset \widetilde{S}(B),$$

and since S_j is a similarity with dilation ratio r_j, we have $L_j \equiv S_j - S_j(0)$ is a rotation and dilation of ratio r_j and thus

$$|\zeta - \eta| = |S_j(x) - S_j(y)| = |L_j(x - y)| \leq r_j |x - y| < r(d(A, B) + \varepsilon),$$

which shows that $\zeta \in \left(\widetilde{S}(B)\right)_{r(d(A,B)+\varepsilon)}$, i.e. (3.8) holds. This completes the proof of the contractive inequality (3.7).

Now let $B = B_R$ be the closed ball as above. The closed ball B is compact by Theorem 11. Set

$$\begin{aligned} F_1 &= \widetilde{S}(B), \\ F_2 &= \widetilde{S}(F_1) = \widetilde{S}^2(B), \\ F_3 &= \widetilde{S}(F_2) = \widetilde{S}^3(B), \\ &\vdots \\ F_k &= \widetilde{S}(F_{k-1}) = \widetilde{S}^k(B), \\ &\vdots \end{aligned}$$

and note that each F_k is a nonempty compact subset of the closed ball B. Indeed, since a similarity maps closed balls to closed balls, each F_k is actually a finite union of closed balls, hence closed by Proposition 7 (4). Moreover, by (3.5), (3.6) and induction we have

$$\begin{aligned} F_1 &= \widetilde{S}(B) \subset B, \\ F_2 &= \widetilde{S}(F_1) \subset \widetilde{S}(B) = F_1, \\ F_3 &= \widetilde{S}(F_2) \subset \widetilde{S}(F_1) = F_2, \\ &\vdots \\ F_k &= \widetilde{S}(F_{k-1}) \subset \widetilde{S}(F_{k-2}) = F_{k-1}, \\ &\vdots \end{aligned}$$

and so the sequence of nonempty compact sets $\{F_k\}_{k=1}^{\infty}$ is nonincreasing. By Corollary 5 we conclude that

$$F \equiv \bigcap_{k=1}^{\infty} F_k$$

is nonempty and compact. Applying $\widetilde{S}$ to F we claim that

$$\widetilde{S}(F) = \bigcap_{k=1}^{\infty} \widetilde{S}(F_k) = \bigcap_{k=1}^{\infty} F_{k+1} = \bigcap_{k=2}^{\infty} F_k = F, \tag{3.9}$$

which proves the existence of a self-similar set satisfying (3.2). The only equality requiring proof in (3.9) is the first. If $\zeta \in \widetilde{S}(F)$ then there is some j and $x \in F$ such that $\zeta = S_j(x)$. Since $F \subset F_k$ we get $\zeta \in S_j(F_k) \subset \widetilde{S}(F_k)$ for all $k \geq 1$, which shows that $\widetilde{S}(F) \subset \bigcap_{k=1}^{\infty} \widetilde{S}(F_k)$. Conversely, suppose that $\zeta \in \bigcap_{k=1}^{\infty} \widetilde{S}(F_k)$. Then for each k there is some j_k and $x_k \in F_k$ such that $\zeta = S_{j_k}(x_k)$. Now there is some j that occurs infinitely often among the j_k. With such a j fixed let $A = \{k \in \mathbb{N} : j_k = j\}$. Then $\zeta = S_j(x_k)$ for all $k \in A$ and since S_j is one-to-one we conclude that $x \equiv S_j^{-1}\zeta$ satisfies $x = x_k \in F_k$ for all $k \in A$. Since A is infinite and $\{F_k\}_{k=1}^{\infty}$ is nonincreasing, we see that $x \in \bigcap_{k=1}^{\infty} F_k = F$. Thus $\zeta = S_j(x) \in S_j(F) \subset \widetilde{S}(F)$, which proves $\bigcap_{k=1}^{\infty} \widetilde{S}(F_k) \subset \widetilde{S}(F)$.

Finally, we use the contractive inequality (3.7) to prove uniqueness. Indeed, suppose that G is another nonempty compact set satisfying $\widetilde{S}(G) = G$. Then from (3.7) we have

$$0 \leq d(F,G) = d\left(\widetilde{S}(F), \widetilde{S}(G)\right) \leq r d(F,G),$$

which implies $d(F,G) = 0$ since $0 < r < 1$. It follows that $F = G$ since d is a metric.

3.4. A paradoxical set. A similarity S with dilation ratio $r = 1$ is said to be a *rigid motion*, i.e. S is a rigid motion if it is a composition of a translation and a rotation. (Note that the very first step in the proof of Theorem 13 breaks down for a rigid motion.) A subset E of Euclidean space $\mathbb{R}^n$ is said to be *paradoxical* if there are subsets A_i, B_j of E, $1 \leq i \leq \ell$, $1 \leq j \leq m$, and rigid motions S_i, T_j such that

$$\begin{aligned} E &= \left(\dot{\cup}_{i=1}^{\ell} A_i\right) \dot{\cup} \left(\dot{\cup}_{j=1}^{m} B_j\right), \\ E &= \dot{\cup}_{i=1}^{\ell} S_i A_i = \dot{\cup}_{j=1}^{m} T_j B_j. \end{aligned} \tag{3.10}$$

The notation $\dot{\cup}$ asserts that the indicated union is pairwise disjoint. The paradox here is that (3.10) says that E can be decomposed into finitely many pairwise disjoint pieces, which can then be rearranged by rigid motions into *two* copies of E.

A famous paradox of Banach and Tarski asserts that the unit ball $B = B(0,1)$ in $\mathbb{R}^3$ is paradoxical, and moreover needs only 5 pieces to witness the paradox: there is a decomposition

$$B = B_1 \dot{\cup} B_2 \dot{\cup} B_3 \dot{\cup} B_4 \dot{\cup} B_5,$$

of B into five pairwise disjoint sets, and there are rigid motions $S_1, ..., S_5$ such that

$$\begin{aligned} B &= S_1(B_1)\,\dot{\cup} S_2(B_2) \\ &= S_3(B_3)\,\dot{\cup} S_4(B_4)\,\dot{\cup} S_5(B_5). \end{aligned}$$

In other words we can break the ball B into five pieces and then using rigid motions, we can rearrange the first two pieces into B itself and rearrange the other three pieces into a separate copy of B. This creates two distinct balls of radius one out of a single ball of radius one using only a decomposition into five pieces and rigid motions. In fact the paradox can be extended to show that if A and B are *any* two bounded subsets of $\mathbb{R}^3$, each containing some ball, then A can be broken into finitely many pieces that can be rearranged to form B. However, the Banach-Tarski paradox requires the axiom of choice. See e.g. [**7**] for details.

It is somewhat surprising that there exists a paradoxical subset E of the plane $\mathbb{R}^2 = \mathbb{C}$ that does *not* require the axiom of choice for its construction, namely the Sierpiński-Mazurkiewicz Paradox: let $e^{i\theta}$ be a transcendental complex number and define sets of complex numbers by

$$\begin{aligned} E &= \left\{ x = \sum_{n=0}^{\infty} x_n e^{in\theta} \in \mathbb{C} : x_n \in \mathbb{Z}_+ \text{ and } x_n = 0 \text{ for all but finitely many } n \right\}, \\ E_1 &= \{x \in E : x_0 = 0\}, \\ E_2 &= \{x \in E : x_0 > 0\}. \end{aligned}$$

Then $E = E_1 \dot{\cup} E_2 = e^{-i\theta} E_1 = E_2 - 1$. Thus E satisfies the replication formula (3.1) using only *rigid motions* with $k = 1$ and $m = 2$,

$$E = \left(e^{i\theta} E\right) \dot{\cup} (E + 1),$$

and so is paradoxical. The set E has fractal dimension $\frac{\ln m}{\ln k} = \frac{\ln 2}{\ln 1} = \frac{\ln 2}{0} = \infty$, while on the other hand, E is a countable subset of the complex plane.

CHAPTER 4

Sequences and Series

Our main focus in this chapter will be on sequences $\{s_n\}_{n=1}^{\infty}$ whose terms s_n are numbers, either rational, real or complex, i.e. on functions from the natural numbers $\mathbb{N}$ to either $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. A key definition is that of *limit* of such a sequence.

DEFINITION 16. *A complex number L is the* limit *of a complex-valued sequence $\{s_n\}_{n=1}^{\infty}$ provided that for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ (depending on ε) such that*

$$|s_n - L| < \varepsilon, \qquad \textit{for all } n \geq N. \tag{0.11}$$

In this case we write

$$\lim_{n\to\infty} s_n = L.$$

Of course this definition applies equally well to the subsets $\mathbb{Q}$ and $\mathbb{R}$ of $\mathbb{C}$. It turns out that the Least Upper Bound Property of the real numbers $\mathbb{R}$ plays a crucial role in the theory of limits, both in $\mathbb{R}$ and in the complex numbers $\mathbb{C}$. For example, if $\{s_n\}_{n=1}^{\infty}$ is a *nondecreasing* sequence of real numbers, i.e.

$$s_{n+1} \geq s_n \qquad \text{for all } n \in \mathbb{N},$$

that is *bounded above*, i.e. there is a real number M such that

$$s_n \leq M \qquad \text{for all } n \in \mathbb{N},$$

then the limit of the sequence $\{s_n\}_{n=1}^{\infty}$ exists, and is given by

$$\lim_{n\to\infty} s_n = \sup\{s_n : n \geq 1\},$$

where in taking the supremum we are viewing $\{s_n : n \geq 1\}$ as a set of real numbers, rather than as the real-valued function on the natural numbers $\mathbb{N}$ that is denoted by $\{s_n\}_{n=1}^{\infty}$.

To see this, let $E = \{s_n : n \geq 1\}$ and $\alpha = \sup E$. Given $\varepsilon > 0$, the number $\alpha - \varepsilon$ is not an upper bound for E and it follows that there is a term s_N such that

$$\alpha - \varepsilon < s_N.$$

Since the sequence $\{s_n\}_{n=1}^{\infty}$ is nondecreasing and bounded above by α, we have

$$\alpha - \varepsilon < s_N \leq s_n \leq \alpha$$

for all $n \geq N$. But this implies that (0.11) holds with $L = \alpha$. We have thus proved the following lemma.

LEMMA 8. *If $\{s_n\}_{n=1}^{\infty}$ is a nondecreasing sequence of real numbers that is bounded above, then $\lim_{n\to\infty} s_n = \sup\{s_n\}_{n=1}^{\infty}$. Similarly, if $\{s_n\}_{n=1}^{\infty}$ is a nonincreasing sequence of real numbers that is bounded below, then $\lim_{n\to\infty} s_n = \inf\{s_n\}_{n=1}^{\infty}$.*

However, later applications of analysis to existence of fractals and solutions to differential equations, will require the notion of sequences of functions in certain metric spaces. Thus we will now develop the critical concepts of limit, subsequence and Cauchy sequence in the broader context of a general metric space.

1. Sequences in a metric space

Recall from Definition 8 that a sequence $\{s_n\}_{n=1}^{\infty}$ is a function f defined on the natural numbers $\mathbb{N}$ with $f(n) = s_n$ for all $n \in \mathbb{N}$. We begin with the general definition of limit.

Definition 17. *Let (X, d) be a metric space. An element L in X is the* limit *of an X-valued sequence $\{s_n\}_{n=1}^{\infty}$ provided that for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ (depending on ε) such that*

$$d(s_n, L) < \varepsilon, \qquad \text{for all } n \geq N. \tag{1.1}$$

In this case we write

$$\lim_{n\to\infty} s_n = L,$$

and say that the sequence $\{s_n\}_{n=1}^{\infty}$ converges to L; otherwise we say $\{s_n\}_{n=1}^{\infty}$ diverges.

Note that limits, if they exist, are unique! Indeed, if both L and L' in X satisfy (1.1), then given $\varepsilon > 0$, there is N so that (1.1) holds for *both* L and L'. Thus the triangle inequality yields

$$0 \leq d(L, L') \leq d(L, s_N) + d(s_N, L') < \varepsilon + \varepsilon = 2\varepsilon.$$

Since ε can be made arbitrarily small, it follows that $d(L, L') = 0$, hence $L = L'$. Here are three more properties of limits that follow easily from Definition 17.

Proposition 10. *Let $\{s_n\}_{n=1}^{\infty}$ be a sequence in a metric space (X, d).*

1. *$\lim_{n\to\infty} s_n = L \in X$ if and only if every ball $B(L, r)$, $r > 0$, contains all of the terms s_n except for finitely many $n \in \mathbb{N}$.*
2. *$\lim_{n\to\infty} s_n = L \in X$ implies that the set $\{s_n\}_{n=1}^{\infty}$ is bounded.*
3. *If $E \subset X$ and if $p \in X$ is a limit point of E, then there is a sequence $\{s_n\}_{n=1}^{\infty}$ in E such that $p = \lim_{n\to\infty} s_n$.*

Proof: (1) Suppose that $\lim_{n\to\infty} s_n = L \in X$ and that $r > 0$. Then there is N such that (1.1) holds with $\varepsilon = r$. Thus $s_n \in B(L, r)$ for all $n \geq N$, and so the only terms s_n not contained in $B(L, r)$ are among the finitely many terms $s_1, \ldots, s_{N-1}$. Conversely, suppose that every ball $B(L, r)$, $r > 0$, contains all of the terms s_n except for finitely many $n \in \mathbb{N}$. Let $\varepsilon > 0$ be given. Then $B(L, \varepsilon)$ contains all but finitely many of the terms s_n. Let M be the largest subscript among these finitely many terms s_n. Then with $N = M + 1$ we have $s_n \in B(L, \varepsilon)$ for all $n \geq N$, which is (1.1). **Note**: Uniqueness of limits follows from (1) as well. Why?

(2) There is N such (1.1) holds with $\varepsilon = 1$. Now set

$$r = \max\{1, d(L, s_1), d(L, s_2), \ldots, d(L, s_N)\}.$$

Then $d(L, s_n) \leq r < r + 1$ for all $n \in \mathbb{N}$ and it follows $\{s_n\}_{n=1}^{\infty} \subset B(L, r + 1)$, i.e. $\{s_n\}_{n=1}^{\infty}$ is bounded in X.

(3) For each $n \in \mathbb{N}$ choose $s_n \in B'\left(p, \frac{1}{n}\right) \cap E$. We claim that $\lim_{n\to\infty} s_n = p$. Indeed, given $\varepsilon > 0$, choose $N \geq \frac{1}{\varepsilon}$. Then for $n \geq N$ we have

$$d(p, s_n) < \frac{1}{n} \leq \frac{1}{N} \leq \varepsilon,$$

as required.

1.1. Subsequences. A key construct associated with a sequence $s = \{s_n\}_{n=1}^{\infty}$ is that of a *subsequence*. A subsequence is defined by viewing s as a map defined on the natural numbers $\mathbb{N}$ and composing it with a *strictly increasing* map $k \to n_k$ from $\mathbb{N}$ to $\mathbb{N}$, to get a map

$$k \to n_k \to s_{n_k}$$

defined on $\mathbb{N}$. In other words we consider a sequence $\{n_k\}_{k=1}^{\infty}$ of strictly increasing positive integers and define the 'composition of sequences' $\{s_{n_k}\}_{k=1}^{\infty}$ to be a *subsequence* of $\{s_n\}_{n=1}^{\infty}$. For example let $\{s_n\}_{n=1}^{\infty}$ be the sequence

$$\{s_n\}_{n=1}^{\infty} = \left\{\frac{\sqrt{n}-1}{\sqrt{n}+1}\right\}_{n=1}^{\infty} = \left\{0, \frac{\sqrt{2}-1}{\sqrt{2}+1}, \frac{\sqrt{3}-1}{\sqrt{3}+1}, \frac{2-1}{2+1}, \frac{\sqrt{5}-1}{\sqrt{5}+1}, ...\right\}.$$

If we take $\{n_k\}_{k=1}^{\infty} = \left\{k^2\right\}_{k=1}^{\infty}$ to be the increasing sequence of square numbers, the corresponding subsequence$\{s_{n_k}\}_{n_k=1}^{\infty}$ of $\{s_n\}_{n=1}^{\infty}$ is given by

$$\{s_{n_k}\}_{n_k=1}^{\infty} = \left\{\frac{\sqrt{k^2}-1}{\sqrt{k^2}+1}\right\}_{n=1}^{\infty} = \left\{\frac{k-1}{k+1}\right\}_{n=1}^{\infty} = \left\{0, \frac{2-1}{2+1}, \frac{3-1}{3+1}, ...\right\}.$$

Note that the terms $\frac{k-1}{k+1}$ in $\{s_{n_k}\}_{n_k=1}^{\infty}$ appear in increasing order among the terms $\frac{\sqrt{n}-1}{\sqrt{n}+1}$ of $\{s_n\}_{n=1}^{\infty}$.

EXERCISE 4. *A sequence $s = \{s_n\}_{n=1}^{\infty}$ converges to L* if and only if *every subsequence $\{s_{n_k}\}_{n_k=1}^{\infty}$ of s converges to L. This is an easy consequence of definition chasing.*

THEOREM 14. *Suppose that $s = \{s_n\}_{n=1}^{\infty}$ is a sequence in a metric space (X, d).*

(1) *If X is* compact*, then some subsequence of s converges to a point in X.*
(2) *If X is Euclidean space $\mathbb{R}^n$ and s is* bounded*, then some subsequence of s converges to a point in $\mathbb{R}^n$.*

We often abbreviate the expression "then some subsequence of s converges to a point in X" to simply "s has a convergent subsequence in X".

Proof: (1) Let E be the set of points $\{s_n : n \in \mathbb{N}\}$. If E is finite, then one of its members, say p, occurs infinitely often in the sequence $s = \{s_1, s_2, s_3, ...\}$. Thus there is a strictly increasing sequence of positive integers

$$n_1 < n_2 < n_3 < ... < n_k < ...$$

such that

$$p = s_{n_1} = s_{n_2} = s_{n_3} = ... = s_{n_k} = ...$$

for all $k \geq 1$. The subsequence $\{s_{n_k}\}_{k=1}^{\infty} = \{p, p, p, ...\}$ clearly converges to $p \in X$.

On the other hand, if E is infinite, then since X is compact, Theorem 7 shows that E has a limit point $p \in X$.

REMARK 10. *Proposition 10 (3) shows there is a sequence $\{t_n\}_{n=1}^{\infty}$ in E that converges to p, but this sequence need* not *be a subsequence of $\{s_n\}_{n=1}^{\infty}$.*

So instead of using Proposition 10 (3), we construct a subsequence of s converging to p as follows: pick n_1 such that $d(p, s_{n_1}) < 1$. Then since $B'(p,1)$ contains *infinitely* many points from E, there is $n_2 > n_1$ such that $d(p, s_{n_2}) < \frac{1}{2}$. Continuing in this way we obtain for every $k \geq 1$ a positive integer n_k such that

$$n_1 < n_2 < ... < n_k < n_{k+1} < ...$$

and $d(p, s_{n_k}) < \frac{1}{k}$ for all $k \geq 1$. Thus the subsequence $\{s_{n_k}\}_{k=1}^{\infty}$ converges to p.

(2) Since $E = \{s_n : n \in \mathbb{N}\}$ is bounded, its closure $\overline{E}$ is closed and bounded in $\mathbb{R}^n$ (bounded since if $E \subset B(x,R)$ then $\overline{E} \subset \overline{B(x,R)} \subset B(x, R+1)$). By Theorem 11 it follows that $\overline{E}$ is compact. Now we can apply part (1) of the theorem, which we just finished proving, with $X = \overline{E}$. This completes the proof of part (2).

In Lemma 3 we proved that the derived set E' of a set E is always closed. We have the following variant for sequences $s = \{s_n\}_{n=1}^{\infty}$ in a metric space X. A point $p \in X$ is said to be a *subsequential limit* of s if $\lim_{k\to\infty} s_{n_k} = p$ for some subsequence $\{s_{n_k}\}_{k=1}^{\infty}$ of s.

THEOREM 15. *The subsequential limits of a sequence $s = \{s_n\}_{n=1}^{\infty}$ in a metric space X form a* closed *subset of X.*

Proof: Let E^* be the set of subsequential limits of s, i.e. all limits of subsequences of s. Suppose that $z \in (E^*)'$. We must show that $z \in E^*$. Now there is $y_1 \in B'\left(z, \frac{1}{2}\right) \cap E^*$ and also n_1 such that $d(y_1, s_{n_1}) < \frac{1}{2}$. Thus we have

$$d(z, s_{n_1}) \leq d(z, y_1) + d(y_1, s_{n_1}) < \frac{1}{2} + \frac{1}{2} = 1.$$

In similar fashion we can choose $n_2 > n_1$ such that $d(z, s_{n_2}) < \frac{1}{2}$. Continuing we can choose $n_1 < n_2 < n_3 < ...$ so that

$$d(z, s_{n_k}) < \frac{1}{k}, \qquad k \geq 1.$$

This shows that the subsequence $\{s_{n_k}\}_{k=1}^{\infty}$ of s converges to z, and hence $z \in E^*$ as required.

1.2. Cauchy sequences. Sequences $\{s_n\}_{n=1}^{\infty}$ of rational numbers $\mathbb{Q}$ can diverge for two qualitatively quite different reasons:

(1) The sequences $\{n\}_{n=1}^{\infty}$ and $\{(-1)^n\}_{n=1}^{\infty}$ fail to converge because the terms s_m and s_n don't even get close to each other, much less close to a limiting value L, as m and n get large.
(2) The sequence $\{s_n\}_{n=1}^{\infty} = \{1.4, 1.41, 1.414, 1.4142, ...\}$ of decimal approximations to the real number $\sqrt{2}$ has no limit in $\mathbb{Q}$ because the rational numbers have a 'gap' where $\sqrt{2}$ ought to be - this despite the fact that $|s_m - s_n| \leq \frac{1}{10^m}$ for all $m < n$, which shows that the terms s_m and s_n get rapidly close to each other as m and n get large.

The first type of divergence above occurs for natural reasons, but the second type of divergence occurs only because of a defect in the metric space $\mathbb{Q}$. The real numbers $\mathbb{R}$ do not share this defect, and Cantor's construction of the real numbers using cuts keyed on the fact that the defect in $\mathbb{Q}$ was a *gap in the order*. We now wish to investigate to what extent this defect can be realized in the *metric space structure* associated with $\mathbb{Q}$ and $\mathbb{R}$, rather than in the *order structure*. As a byproduct of this investigation, we will be led to Weierstrass' construction of the real numbers using *Cauchy sequences* of rational numbers. Our first definition captures the notion of a sequence $\{s_n\}_{n=1}^{\infty}$ of the second type above in which the terms s_m and s_n get close to each other as m and n get large, and so 'ought' to have a limit in a 'nondefective' metric space.

DEFINITION 18. *Let (X, d) be a metric space. A sequence $\{s_n\}_{n=1}^{\infty}$ in X is a* Cauchy *sequence if for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that*

$$d(s_m, s_n) < \varepsilon, \qquad \text{for all } m, n \geq N. \tag{1.2}$$

LEMMA 9. *Convergent sequences in a metric space are Cauchy sequences.*

Proof: Suppose $\{s_n\}_{n=1}^{\infty}$ is a convergent sequence in a metric space (X, d), i.e. $\lim_{n\infty} s_n = L$ for some $L \in X$. Let $\varepsilon > 0$ be given. Choose N as in Definition 17 so that $d(s_n, L) < \frac{\varepsilon}{2}$ for all $n \geq N$. Then if $m, n \geq N$, the triangle inequality yields

$$d(s_m, s_n) \leq d(s_m, L) + d(L, s_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

There is a partial converse to this lemma.

LEMMA 10. *Let $s = \{s_n\}_{n=1}^{\infty}$ be a Cauchy sequence in a metric space X. Then s converges* if and only if *it has a convergent subsequence in X.*

Proof: If $\{s_n\}_{n=1}^{\infty}$ converges in a metric space X to a limit L, then *every* subsequence converges to L as well. Conversely suppose that $s = \{s_n\}_{n=1}^{\infty}$ is a Cauchy sequence in X and that $\lim_{k\to\infty} s_{n_k} = L \in X$ for some subsequence $\{s_{n_k}\}_{k=1}^{\infty}$. Given $\varepsilon > 0$ the Cauchy criterion (1.2) yields N so that

$$d(s_m, s_n) < \frac{\varepsilon}{2}, \qquad m, n \geq N,$$

and then the definition of limit yields K satisfying

$$d(s_{n_k}, L) < \frac{\varepsilon}{2}, \qquad k \geq K.$$

We may also take K so large that $n_K \geq N$. Then for $n \geq N$ we have

$$d(s_n, L) \leq d(s_n, s_{n_K}) + d(s_{n_K}, L) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which shows that $\lim_{n\to\infty} s_n = L$.

Now comes our definition of a 'nondefective' metric space, which we call *complete*.

DEFINITION 19. *A metric space X is* complete *if every Cauchy sequence in X converges to a point in X.*

Roughly speaking, a complete metric space X has the property that any sequence which *ought* to converge, i.e. one that satisfies the Cauchy criterion, actually *does* converge in X. In a complete metric space, the condition that for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ satisfying (1.2), is often called the *Cauchy criterion* for convergence of the sequence $\{s_n\}_{n=1}^{\infty}$.

The crucial difference between the rational and real numbers can now be expressed in metric terms: the space $\mathbb{Q}$ is *not* complete whereas the space $\mathbb{R}$ *is* complete. In order to prove our theorem on completeness it is convenient to introduce the concept of *diameter* of a set. If A is a subset of real numbers, we extend the definition of $\sup A$ to sets that are *not* bounded above by defining

$$\sup A = \infty, \qquad \text{if } A \text{ is not bounded above.}$$

DEFINITION 20. *If E is a subset of a metric space (X, d), we define the* diameter *of E to be*

$$diam\,(E) = \sup \{d\,(x, y) : x, y \in E\}\,.$$

The connection with Cauchy sequences is this. Suppose $s = \{s_n\}_{n=1}^{\infty}$ is a sequence in a metric space (X, d). Let $T_N = \{s_n : n \geq N\}$ be the set of points in the tail of the sequence from N on. Then s is a Cauchy sequence if and only if

$$diam\,(T_N) \to 0 \text{ as } N \to \infty. \tag{1.3}$$

The reader can easily verify this by chasing definitions.

LEMMA 11. $diam\,(E) = diam\left(\overline{E}\right)$.

Proof: Clearly $diam\,(E) \leq diam\left(\overline{E}\right)$ holds since $E \subset \overline{E}$. Conversely pick $\varepsilon > 0$ and two points p, q in $\overline{E}$. There are points $x, y \in E$ such that $d\,(p, x) < \frac{\varepsilon}{2}$ and $d\,(q, y) < \frac{\varepsilon}{2}$. Thus we have

$$d\,(p, q) \leq d\,(p, x) + d\,(x, y) + d\,(y, q) \leq \frac{\varepsilon}{2} + diam\,(E) + \frac{\varepsilon}{2} = diam\,(E) + \varepsilon,$$

even in the case that $diam\,(E) = \infty$. Now take the infimum over $\varepsilon > 0$ to obtain $d\,(p, q) \leq diam\,(E)$ for all $p, q \in E$, and then take the supremum over all such p, q to obtain $diam\left(\overline{E}\right) \leq diam\,(E)$ as required.

THEOREM 16. *Let X be a metric space.*

(1) *If X is compact, then X is complete.*
(2) *Euclidean space $\mathbb{R}^n$ is complete.*

Proof: (1) Suppose that $\{s_n\}_{n=1}^{\infty}$ is a Cauchy sequence in a compact metric space X. Let $T_N = \{s_n : n \geq N\}$ be the set of points in the tail of the sequence from N on. By (1.3) the Cauchy criterion says that $diam\,(T_N) \to 0$ as $N \to \infty$. Lemma 11 then gives

$$diam\left(\overline{T_N}\right) \to 0 \text{ as } N \to \infty. \tag{1.4}$$

Now $\overline{T_N}$ is nonempty and compact for each N, and clearly $\overline{T_{N+1}} \subset \overline{T_N}$ for all N. Corollary 5 thus shows that

$$K = \bigcap_{N=1}^{\infty} \overline{T_N} \neq \emptyset.$$

Since $K \subset \overline{T_N}$, (1.4) gives $diam\,(K) = 0$, from which we conclude that K consists of *exactly* one point, say $L \in X$.

We now claim that $\lim_{n\to\infty} s_n = L$. Indeed, given $\varepsilon > 0$, choose N so large that $diam\left(\overline{T_N}\right) < \varepsilon$. Then for all $n \geq N$ we have that both s_n and L belong to $\overline{T_N}$, and so

$$d\,(s_n, L) \leq diam\left(\overline{T_N}\right) < \varepsilon,$$

as required.

(2) Suppose that $s = \{s_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}^n$. There is N so large that the tail T_N has diameter at most 1. Since there are only finitely many points s_n outside the tail T_N, it follows that the set of points $\{s_n : n \geq 1\}$ in the sequence

is bounded. The closure of this set is also bounded, and thus s is contained in a closed and bounded subset X of $\mathbb{R}^n$. By Theorem 11 the set X is compact and we can now apply part (1) of the theorem proved above.

1.3. Weierstrass' construction of the real numbers. Recall that the gap in the rational numbers where the irrational number $\sqrt{2}$ lives, can be detected either by one of Dedekind's cuts (the cut A in (2.1) of Chapter 1), or by the Cauchy sequence of decimal approximations $\{1.4, 1.41, 1.414, ...\}$. While Dedekind used cuts to construct the real numbers, Weierstrass instead used such Cauchy sequences in $\mathbb{Q}$ to construct the real numbers by filling in these gaps in the rationals as follows.

Denote by $\mathcal{C}$ the set of all Cauchy sequences $s = \{s_n\}_{n=1}^{\infty}$ in $\mathbb{Q}$. Define an equivalence relation on $\mathcal{C}$ by $s \sim t$ if the intertwined sequence

$$\{s_1, t_1, s_2, t_2, s_3, t_3, ...\}$$

is also a Cauchy sequence (intuitively this says that the limits that s and t ought to have should coincide). Once we have proved this relation is indeed an equivalence relation, then we can define the equivalence class $[s]$ of a Cauchy sequence s in $\mathcal{C}$, and we can define the real numbers $\mathbb{R}$ to be the set of equivalence classes:

$$\mathbb{R} = \{[s] : s \in \mathcal{C}\}.$$

At this point the construction becomes as tedious as that of Dedekind, and we omit the details, only mentioning that one defines the *sum* of two classes $[s]$ and $[t]$ where $s, t \in \mathcal{C}$, by proving that the sequence $s + t = \{s_n + t_n\}_{n=1}^{\infty}$ is Cauchy, and then defining

$$[s] + [t] \equiv [s + t].$$

It is a long process to define the remaining relations and verify that $\mathbb{R}$ satisfies the axioms of an ordered field with the least upper bound property.

This method of Weierstrass for constructing the real numbers has an advantage the method of Dedekind lacks. Namely it can be used to construct an extension of an *arbitrary* metric space X to a (usually larger) space $\widehat{X}$ that is complete, and called the completion of X. More precisely, but without much detail, define $\widehat{X}$ to be the set of equivalence classes $[s]$ in the set $\mathcal{C}$ of Cauchy sequences s in X, where $s \sim t$ if $\{s_1, t_1, s_2, t_2, ...\}$ is Cauchy in X. Define a function $\widehat{d}$ on $\widehat{X} \times \widehat{X}$ by

$$\widehat{d}([s], [t]) = \lim_{n \to \infty} d(s_n, t_n).$$

After showing that the limit above exists, and that $\left(\widehat{X}, \widehat{d}\right)$ satisfies the axioms for a metric space, one can prove that the space $\left(\widehat{X}, \widehat{d}\right)$ is *complete*. We can view X as a subspace of $\widehat{X}$ via the map that sends x in X to the equivalence class containing the constant Cauchy sequence $\{x, x, x, ...\}$. One can verify that this map is an isometry, and moreover that under this identification of X with a subspace of $\widehat{X}$, the set X is dense in $\widehat{X}$. This shows that $\widehat{X}$ is, up to an isometry, the smallest complete space containing X, and this is the reason that $\widehat{X}$ is called *the completion* of X.

On the other hand, the idea of a Dedekind cut can only be used to construct an extension of a linearly ordered set to one with the least upper bound property, a concept that has not been nearly so useful in applications of analysis as is the concept of a complete metric space. For example, the next subsection describes one of the most useful results in the theory of abstract metric spaces, one that can be used to simplify the ideas behind the proof of Theorem 13, and to prove many existence theorems for differential equations, as we illustrate in a later chapter.

1.4. A contraction lemma. It is possible to recast the proof of Theorem 13 on the existence and uniqueness of nonempty compact fractals, entirely within the context of the metric space $\mathcal{X}^n$ of compact subsets of $\mathbb{R}^n$ that was introduced above. This is achieved by using the fact that the map $\widetilde{S}: \mathcal{X}^n \to \mathcal{X}^n$ defined in (3.4) is a strict contraction, i.e. satisfies (3.7) for some $0 < r < 1$, defined on the *complete* metric space $\mathcal{X}^n$.

Of course we haven't yet shown that $\mathcal{X}^n$ is complete, and we defer the proof of this to the end of this subsection. The main idea is to use the finite intersection property of compact sets much as we did in the proof of Theorem 13.

Once we know that $\mathcal{X}^n$ is complete, the following Contraction Lemma immediately proves Theorem 13 on the existence and uniqueness of fractals.

LEMMA 12. *Suppose that (X, d) is a complete metric space and that $\varphi : X \to X$ is a strict contraction on X, i.e. there is $0 < r < 1$ such that*

$$d(\varphi(x), \varphi(y)) \leq r d(x, y), \qquad \text{for all } x, y \in X.$$

Then φ has a unique fixed point z in X, i.e. there is $z \in X$ such that $\varphi(z) = z$, and if $w \in X$ is another point satisfying $\varphi(w) = w$, then $z = w$.

Proof: The uniqueness assertion is immediate from

$$0 \leq d(z, w) = d(\varphi(z), \varphi(w)) \leq r d(z, w),$$

since $0 < r < 1$. To establish the existence assertion, pick *any* point $s_0 \in X$. Consider the sequence of iterates $\{s_n\}_{n=1}^\infty$ given by

$$\begin{aligned}
s_1 &= \varphi(s_0), \\
s_2 &= \varphi(s_1) = \varphi(\varphi(s_0)) = \varphi^2(s_0), \\
s_3 &= \varphi(s_2) = \varphi\left(\varphi^2(s_0)\right) = \varphi^3(s_0), \\
&\vdots \\
s_n &= \varphi(s_{n-1}) = \varphi\left(\varphi^{n-1}(s_0)\right) = \varphi^n(s_0), \\
&\vdots
\end{aligned}$$

We claim that the sequence $\{s_n\}_{n=1}^\infty$ is Cauchy. To see this first note that

$$d(s_k, s_{k+1}) = d(\varphi(s_{k-1}), \varphi(s_k)) \leq r d(s_{k-1}, s_k), \qquad k \geq 1,$$

and then use induction to prove that

$$d(s_\ell, s_{\ell+1}) \leq r d(s_{\ell-1}, s_\ell) \leq r^2 d(s_{\ell-2}, s_{\ell-1}) \leq \ldots \leq r^\ell d(s_0, s_1).$$

Now for $m < n$, the triangle inequality yields

$$\begin{aligned} d(s_m, s_n) &\leq d(s_m, s_{m+1}) + d(s_{m+1}, s_{m+2}) + \ldots + d(s_{n-1}, s_n) \\ &= \sum_{j=0}^{n-m} d(s_{m+j}, s_{m+j+1}) \\ &\leq \sum_{j=0}^{n-m} r^{m+j} d(s_0, s_1) \\ &< \frac{r^m}{1-r} d(s_0, \varphi(s_0)). \end{aligned}$$

Thus given $\varepsilon > 0$, if we choose N so large that $\frac{r^N}{1-r} d(s_0, \varphi(s_0)) < \varepsilon$, then we have $d(s_m, s_n) < \varepsilon$ for all $m, n \geq N$, which proves that $\{s_n\}_{n=1}^{\infty}$ is Cauchy.

Now we use the important hypothesis that X is complete. Thus $\{s_n\}_{n=1}^{\infty}$ is convergent and there is a limit

$$z = \lim_{n\to\infty} s_n \in X.$$

The triangle inequality gives

$$\begin{aligned} d(\varphi(z), z) &\leq d(\varphi(z), \varphi(s_n)) + d(\varphi(s_n), s_{n+1}) + d(s_{n+1}, z) \\ &\leq r d(z, s_n) + 0 + d(z, s_{n+1}) \\ &\leq d(z, s_n) + d(z, s_{n+1}), \end{aligned}$$

which tends to 0 as $n \to \infty$. It follows that $d(\varphi(z), z) = 0$ and hence $\varphi(z) = z$.

LEMMA 13. *The metric space $\mathcal{X}^n$ is complete.*

Proof: Suppose that $\{K_j\}_{j=1}^{\infty}$ is a Cauchy sequence in $\mathcal{X}^n$. For each $\ell \geq 1$ there is by the definition (3.3) of the metric in $\mathcal{X}^n$ together with the Cauchy criterion (1.2), a positive integer j_ℓ such that

$$K_j \subset (K_{j_\ell})_{\frac{1}{2^{\ell+1}}} \text{ and } K_{j_\ell} \subset (K_j)_{\frac{1}{2^{\ell+1}}}, \qquad \text{for all } j \geq j_\ell,$$

and moreover we can choose the j_ℓ to be strictly increasing, i.e. $j_\ell < j_{\ell+1}$ for all $\ell \geq 1$. Using $j = j_{\ell+1} > j_\ell$ we then also have the following inequalities:

$$\left(K_{j_{\ell+1}}\right)_{\frac{1}{2^{\ell+1}}} \subset \left((K_{j_\ell})_{\frac{1}{2^{\ell+1}}}\right)_{\frac{1}{2^{\ell+1}}} \subset (K_{j_\ell})_{\frac{1}{2^\ell}}, \qquad \text{for all } \ell \geq 1.$$

Thus the sequence of closed bounded nonempty sets

$$\left\{\overline{(K_{j_\ell})_{\frac{1}{2^\ell}}}\right\}_{\ell=1}^{\infty}$$

is nonincreasing, and by Theorem 11 consists of compact sets. By Corollary 6 we then conclude that

$$K = \bigcap_{\ell=1}^{\infty} \overline{(K_{j_\ell})_{\frac{1}{2^\ell}}}$$

is a nonempty compact set, so $K \in \mathcal{X}^n$.

We now claim that

$$\lim_{j\to\infty} K_j = K.$$

Since $\{K_j\}_{j=1}^{\infty}$ is Cauchy it suffices by Lemma 10 to prove that

$$\lim_{\ell\to\infty} K_{j_\ell} = K.$$

Let $\delta > 0$ be given. We trivially have

$$K \subset \overline{(K_{j_\ell})_{\frac{1}{2^\ell}}} \subset (K_{j_\ell})_{\frac{1}{2^{\ell-1}}} \subset (K_{j_\ell})_\delta \tag{1.5}$$

for ℓ so large that $2^{\ell-1} > \frac{1}{\delta}$. In the other direction $\left\{\left(\overline{(K_{j_\ell})_{\frac{1}{2^\ell}}}\right)^c\right\}_{\ell=1}^\infty$ is an open cover of the compact set $K_{j_1} \cap K_\delta^c$, and if $\left\{\left(\overline{(K_{j_\ell})_{\frac{1}{2^\ell}}}\right)^c\right\}_{\ell=1}^L$ is a finite subcover, then

$$K_{j_1} \cap (K)_\delta^c \subset \bigcup_{\ell=1}^L \left(\overline{(K_{j_\ell})_{\frac{1}{2^\ell}}}\right)^c,$$

equivalently

$$\overline{(K_{j_L})_{\frac{1}{2^L}}} = \bigcap_{\ell=1}^L \overline{(K_{j_\ell})_{\frac{1}{2^\ell}}} \subset K_{j_1}^c \cup (K)_\delta\,,$$

which implies

$$K_{j_\ell} \subset \overline{(K_{j_\ell})_{\frac{1}{2^\ell}}} \subset (K)_\delta\,, \qquad \text{for all } \ell \geq L. \tag{1.6}$$

Altogether (1.5) and (1.6) show that $d(K, K_{j_\ell}) < \delta$ for ℓ sufficiently large as required.

2. Numerical sequences and series

At the beginning of this chapter we proved in Lemma 8 that bounded monotonic sequences $s = \{s_n\}_{n=1}^\infty$ of real numbers converge, and moreover we identified the limit L as either the least upper bound or the greatest lower bound of the set of terms $E \equiv \{s_n : n \geq 1\}$:

$$\lim_{n\to\infty} s_n = L \equiv \begin{cases} \sup E & \text{if} \quad s \text{ is nondecreasing} \\ \inf E & \text{if} \quad s \text{ is nonincreasing} \end{cases}.$$

Here are some examples of monotonic sequences for which we can further identify the sup or inf as a specific real number:

(1) $\lim_{n\to\infty} \frac{1}{n^p} = 0$ if $p > 0$.
(2) $\lim_{n\to\infty} \sqrt[n]{p} = 1$ if $p > 0$.
(3) $\lim_{n\to\infty} \sqrt[n]{n} = 1$.
(4) $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$ if $p > 0$ and $\alpha \in \mathbb{R}$.
(5) $\lim_{n\to\infty} x^n = 0$ if $-1 < x < 1$.

To prove limit (1), let $\varepsilon > 0$ be given and use the Archimedian property of the real numbers to choose $N > \frac{1}{\sqrt[p]{\varepsilon}}$. Then $0 < \frac{1}{n^p} < \varepsilon$ for all $n \geq N$.

The limit in (2) is trivial if $p = 1$. If $p > 1$ then $\sqrt[n]{p} > 1$ and the binomial theorem for $n \geq 1$ yields

$$\begin{aligned} p &= (\sqrt[n]{p})^n = [1 + (\sqrt[n]{p} - 1)]^n \\ &= 1 + n(\sqrt[n]{p} - 1) + \frac{n(n-1)}{2}(\sqrt[n]{p} - 1)^2 + \dots \\ &> 1 + n(\sqrt[n]{p} - 1)\,, \end{aligned}$$

so that

$$0 < \sqrt[n]{p} - 1 < \frac{p-1}{n}, \qquad n \geq 1,$$

which shows that $\lim_{n\to\infty}\left(\sqrt[n]{p}-1\right)=0$ by limit (1), hence $\lim_{n\to\infty}\sqrt[n]{p}=1$. Finally, if $0<p<1$, apply the result just proved to the number $\frac{1}{p}>1$ to get $\lim_{n\to\infty}\sqrt[n]{\frac{1}{p}}=1$, which gives the desired result upon taking reciprocals.

To see limit (3) we argue as in the proof for (2), but keep the quadratic term in the binomial expansion for $n\geq 2$ instead of the linear term:

$$\begin{aligned} n &= \left(\sqrt[n]{n}\right)^n=\left[1+\left(\sqrt[n]{n}-1\right)\right]^n \\ &= 1+n\left(\sqrt[n]{n}-1\right)+\frac{n\,(n-1)}{2}\left(\sqrt[n]{n}-1\right)^2+... \\ &> 1+\frac{n\,(n-1)}{2}\left(\sqrt[n]{n}-1\right)^2, \end{aligned}$$

so that

$$0<\sqrt[n]{n}-1<\sqrt{\frac{n-1}{\frac{n(n-1)}{2}}}=\sqrt{\frac{2}{n}}=\frac{\sqrt{2}}{\sqrt{n}},\qquad n\geq 2,$$

which shows that $\lim_{n\to\infty}\left(\sqrt[n]{n}-1\right)=0$ by limit (1), hence $\lim_{n\to\infty}\sqrt[n]{n}=1$.

To see (4) let k be a positive integer greater than α. Then for $n>2k$ we have

$$(1+p)^n>\binom{n}{k}p^k=\frac{n\,(n-1)\,...\,(n-k+1)}{k\,(k-1)\,...1}p^k>\left(\frac{n}{2}\right)^k\frac{p^k}{k!},$$

so that

$$0<\frac{n^\alpha}{(1+p)^n}<n^\alpha\left(\frac{2}{n}\right)^k\frac{k!}{p^k}=n^{\alpha-k}\left(\frac{2}{p}\right)^k k!,\qquad n>2k,$$

which shows that $\lim_{n\to\infty}\frac{n^\alpha}{(1+p)^n}=0$ since $\lim_{n\to\infty}n^{\alpha-k}=0$ if $\alpha-k<0$ by limit (1).

Limit (5) is the special case $\alpha=0$ of limit (4).

2.1. Series of complex numbers. Given a sequence $\{a_n\}_{n=1}^\infty$ of complex numbers, we can use the field structure on $\mathbb{C}$ to define the corresponding sequence of partial sums

$$s_N\equiv a_1+a_2+...+a_N=\sum_{n=1}^N a_n$$

for all $N\geq 1$. Now if there were only finitely many nonzero terms a_n in the original sequence, then the sequence of partial sums $\{s_N\}_{N=1}^\infty$ would eventually be constant and that constant would be the sum of the nonzero terms a_n. Thus in this case we have

$$\sum_{n:a_n\neq 0}a_n=\lim_{N\to\infty}s_N.$$

This motivates the definition of the *infinite* sum $\sum_{n=1}^\infty a_n$ as the limit $\lim_{N\to\infty}s_N$ of the partial sums, provided that limit exists.

Definition 21. *Suppose that $\{a_n\}_{n=1}^\infty$ is a sequence of complex numbers. If the sequence of partial sums $\{s_N\}_{N=1}^\infty$, $s_N=\sum_{n=1}^N a_n$, converges to a complex number L, we say that the (infinite) series $\sum_{n=1}^\infty a_n$* converges *to L, and write*

$$\sum_{n=1}^\infty a_n=L.$$

If the sequence of partial sums $\{s_N\}_{N=1}^{\infty}$ *diverges, we say that the series* $\sum_{n=1}^{\infty} a_n$ diverges*.*

Recall that as a metric space, the complex numbers $\mathbb{C}$ are isomorphic to $\mathbb{R}^2$, and hence complete by Theorem 16. The Cauchy criterion thus takes the following form for series:

- The series $\sum_{k=1}^{\infty} a_k$ converges in $\mathbb{C}$ if and only if for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that
$$\left| \sum_{k=m}^{n} a_k \right| < \varepsilon, \qquad \text{for all } m, n \geq N.$$

This is easily seen using the Cauchy criterion for the sequence of partial sums $\{s_N\}_{N=1}^{\infty}$, together with the fact that $s_n - s_{m-1} = \sum_{k=m}^{n} a_k$. Note that this provides a simple necessary condition for convergence $\sum_{n=1}^{\infty} a_n$, namely

$$|a_n| \to 0 \text{ as } n \to \infty. \tag{2.1}$$

The reader is cautioned however, that (2.1) is *not* in general sufficient for convergence of the series $\sum_{n=1}^{\infty} a_n$. For example, if $a_n = \frac{1}{n}$ then (2.1) holds but the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges since the partial sums of order $N = 2^k$ satisfy

$$\begin{aligned} s_N &= \sum_{n=1}^{2^k} \frac{1}{n} = \frac{1}{1} + \left(\frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^{k-1}+1} + \ldots + \frac{1}{2^k}\right) \\ &\geq \frac{1}{1} + \left(\frac{1}{2}\right) + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^k} + \ldots + \frac{1}{2^k}\right) \\ &= 1 + \frac{k}{2}, \end{aligned}$$

which is unbounded, and hence the sequence $\{s_N\}_{N=1}^{\infty}$ cannot converge.

We also note the following *sufficient* condition for the convergence of $\sum_{k=1}^{\infty} a_k$:

$$\sum_{k=1}^{\infty} |a_k| \text{ converges.} \tag{2.2}$$

Indeed, if $\sum_{k=1}^{\infty} |a_k|$ converges, say to $L \geq 0$, then we have

$$\left| \sum_{k=m}^{n} a_k \right| \leq \sum_{k=m}^{n} |a_k| = t_n - t_{m-1}, \tag{2.3}$$

where $t_N = \sum_{k=1}^{n} |a_k|$ is the N^{th} partial sum of the series $\sum_{k=1}^{\infty} |a_k|$. Now (2.2) implies that $\{t_N\}_{N=1}^{\infty}$ satisfies the Cauchy criterion for sequences, and together with (2.3), this proves the Cauchy criterion for the series $\sum_{n=1}^{\infty} a_n$. Thus the series $\sum_{n=1}^{\infty} a_n$ converges. Note that the same argument proves the convergence of $\sum_{n=1}^{\infty} a_n$ if $|a_n| \leq b_n$ for all sufficiently large n where $\sum_{n=1}^{\infty} b_n$ converges. We have just proved the first half of the versatile *Comparison Test*. The second half is a trivial consequence of the first.

THEOREM 17. *Suppose that* $\{a_n\}_{n=1}^{\infty}$ *is a sequence of complex numbers.*

(1) *If* $|a_n| \leq b_n$ *for all sufficiently large* n*, and if* $\sum_{n=1}^{\infty} b_n$ converges*, then so does* $\sum_{n=1}^{\infty} a_n$*.*
(2) *If* $a_n \geq b_n \geq 0$ *for all sufficiently large* n*, and if* $\sum_{n=1}^{\infty} b_n$ diverges*, then so does* $\sum_{n=1}^{\infty} a_n$*.*

Probably the most used fact about series of complex numbers is the geometric series formula.

LEMMA 14. *If* $|z| < 1$*, then*

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}.$$

If $|z| \geq 1$*, then* $\sum_{n=0}^{\infty} z^n$ *diverges.*

Proof: The partial sums are given by $s_N = \sum_{n=0}^{N} z^n = \frac{1-z^{N+1}}{1-z}$ for $N \geq 1$. Now $\left|z^{N+1}\right| = |z|^{N+1} \to 0$ as $N \to \infty$ if $|z| < 1$ by limit (5) in the previous subsection, and so

$$\sum_{n=0}^{\infty} z^n \equiv \lim_{N\to\infty} s_N = \lim_{N\to\infty} \frac{1-z^{N+1}}{1-z} = \frac{1}{1-z}, \qquad \text{for } |z| < 1.$$

If on the other hand, $|z| \geq 1$, then $|z^n| = |z|^n$ does not tend to 0 as $n \to \infty$, and hence the series $\sum_{n=0}^{\infty} z^n$ can't converge by (2.1).

EXAMPLE 7. *The series* $\sum_{n=1}^{\infty} \frac{\sin(n\theta)}{n^n}$ *converges for every real* θ *since* $\left|\frac{\sin(n\theta)}{n^n}\right| \leq \frac{1}{2^n}$ *for all* $n \geq 2$*. Indeed,* $\sum_{n=1}^{\infty} \frac{1}{2^n}$ *converges by Lemma 14, and the comparison test Theorem 17 then shows that* $\sum_{n=1}^{\infty} \frac{\sin(n\theta)}{n^n}$ *converges.*

In order to take advantage of the comparison test as we did in the example above, we must have available a large supply of series $\sum_{n=1}^{\infty} b_n$ with nonnegative terms b_n, for which we *already* know whether or not $\sum_{n=1}^{\infty} b_n$ converges. So we now turn to the investigation of series with nonnegative terms.

2.2. Series of nonnegative terms. Lemma 8 on the convergence of increasing sequences has the following useful reformulation for series with nonnegative terms.

LEMMA 15. *Suppose that* $\sum_{n=1}^{\infty} a_n$ *is a series of* nonnegative *terms* a_n*, and let* $s_N = \sum_{n=1}^{N} a_n$ *be the* N^{th} *partial sum. Then the series* $\sum_{n=1}^{\infty} a_n$ *converges if and only if the sequence of partial sums* $\{s_N\}_{N=1}^{\infty}$ *is bounded.*

Proof: We simply chase the definitions with Lemma 8 as follows. The series $\sum_{n=1}^{\infty} a_n$ converges if and only if the sequence $\{s_N\}_{N=1}^{\infty}$ has a limit. But $s_N - s_{N-1} = a_N \geq 0$ shows that the sequence $\{s_N\}_{N=1}^{\infty}$ is nondecreasing. Thus Lemma 8 shows that $\{s_N\}_{N=1}^{\infty}$ has a limit if and only if the sequence is bounded.

Our first main result in this subsection is the *Cauchy condensation test* that applies to a series $\sum_{n=1}^{\infty} a_n$ of *nonincreasing* positive terms a_n and says that the series

$$a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8 + a_9 + ...$$

converges if and only if the *condensed* series

$$\begin{aligned} & a_1 + (a_2 + a_2) + (a_4 + a_4 + a_4 + a_4) + (a_8 + a_8 + ...a_8) + ... \\ = \quad & a_1 + 2a_2 + 4a_4 + 8a_8 + ... \end{aligned} \tag{2.4}$$

converges. Note that the definition of the condensed series is motivated by regrouping the terms in $\sum_{n=1}^{\infty} a_n$ as

$$a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + (a_8 + a_9 + ... + a_{15}) + ... \tag{2.5}$$

THEOREM 18. *Suppose that* $\sum_{n=1}^{\infty} a_n$ *is a series of* nonincreasing *positive terms* a_n. *Then the series* $\sum_{n=1}^{\infty} a_n$ *converges* if and only if *its condensed series* $\sum_{k=0}^{\infty} 2^k a_{2^k}$ *converges.*

Proof: Let $s_N = \sum_{n=1}^{N} a_n$ be the partial sums of the series $\sum_{n=1}^{\infty} a_n$ and let $t_K = \sum_{k=0}^{K} 2^k a_{2^k}$ be the partial sums of the condensed series $\sum_{k=0}^{\infty} 2^k a_{2^k}$ in the second line of (2.4). Suppose first that $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges. We will use the grouping of terms indicated in (2.5). For $N = 2^{K+1} - 1$ we have

$$(2.6)\quad s_N = \sum_{n=1}^{2^{K+1}-1} a_n = \sum_{k=0}^{K} \left(\sum_{n=2^k}^{2^{k+1}-1} a_n \right) \leq \sum_{k=0}^{K} \left(\sum_{n=2^k}^{2^{k+1}-1} a_{2^k} \right) = \sum_{k=0}^{K} 2^k a_{2^k} = t_K,$$

where the inequality follows from the assumption that the terms a_n are positive and nonincreasing. The convergence of $\sum_{k=0}^{\infty} 2^k a_{2^k}$ shows that the partial sums $\{t_K\}_{K=0}^{\infty}$ are bounded, and (2.6) now shows that the subsequence of partial sums $\{s_{2^{K+1}-1}\}_{K=0}^{\infty}$ is bounded. Since the full sequence of partial sums $\{s_N\}_{N=1}^{\infty}$ is nondecreasing, we conclude that it is bounded as well. Then Lemma 15 shows that the series $\sum_{n=1}^{\infty} a_n$ converges.

Conversely we use an inequality opposite to (2.6) that is suggested by the alternate grouping of terms in the series $\sum_{n=1}^{\infty} a_n$ given by (compare with (2.5)),

$$a_1 + (a_2) + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \ldots$$

The inequality is that for $N = 2^K$ we have

$$\begin{aligned} (2.7)\ s_N &= \sum_{n=1}^{2^K} a_n = a_1 + \sum_{k=1}^{K} \left(\sum_{n=2^{k-1}+1}^{2^k} a_n \right) \geq a_1 + \sum_{k=1}^{K} \left(\sum_{n=2^{k-1}+1}^{2^k} a_{2^k} \right) \\ &= a_1 + \sum_{k=1}^{K} 2^{k-1} a_{2^k} = \frac{1}{2}\left(a_1 + t_K\right), \end{aligned}$$

where again the inequality follows from the assumption that the terms a_n are positive and nonincreasing. If $\sum_{n=1}^{\infty} a_n$ converges, then the sequence of partial sums $\{s_N\}_{N=1}^{\infty}$ is bounded, and (2.7) shows that the sequence of partial sums $\{t_K\}_{K=0}^{\infty}$ is bounded, hence $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges by Lemma 3.4.

COROLLARY 8. *Let* $p \in \mathbb{R}$. *The p-series* $\sum_{n=1}^{\infty} \frac{1}{n^p}$ *converges if and only if* $p > 1$.

Proof: For $p \leq 0$ the series diverges since $\frac{1}{n^p}$ does *not* go to zero as $n \to \infty$. If $p > 0$ then the terms $\frac{1}{n^p}$ are nonincreasing and so the Cauchy condensation test shows that $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if and only if its condensed series

$$\sum_{k=0}^{\infty} 2^k \frac{1}{2^{kp}} = \sum_{k=0}^{\infty} \left(\frac{1}{2^{p-1}} \right)^k$$

converges. But the condensed series is a geometric series and Lemma 14 shows that it converges if and only if $\frac{1}{2^{p-1}} < 1$, i.e. $p > 1$.

The series of reciprocals of factorials,

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{1}{n!} &= \frac{1}{0!}+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+...+\frac{1}{n!}+...\\ &= 1+1+\frac{1}{2\cdot 1}+\frac{1}{3\cdot 2\cdot 1}+...+\frac{1}{n\,(n-1)\,...3\cdot 2\cdot 1}+...\end{aligned}$$

plays a very distinguished role in analysis. First we note that this series converges by the comparison test and the geometric series formula. Indeed,

$$\frac{1}{n!}=\frac{1}{n\,(n-1)\,...3\cdot 2\cdot 1}\leq\frac{1}{2\,(2)\,...2\cdot 2\cdot 1}=\left(\frac{1}{2}\right)^{n-1}$$

for all $n\geq 2$, and

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

by Lemma 14. Thus $\sum_{n=0}^{\infty}\frac{1}{n!}$ converges by Theorem 17 (1), and in fact

$$2<1+1+\sum_{n=2}^{\infty}\frac{1}{n!}=\sum_{n=0}^{\infty}\frac{1}{n!}<1+1+\sum_{n=2}^{\infty}\left(\frac{1}{2}\right)^{n-1}=3.$$

Definition 22. $e\equiv\sum_{n=0}^{\infty}\frac{1}{n!}$.

The series for Euler's number e converges so rapidly that it forces e to be irrational. Indeed, if $s_N=\sum_{n=0}^{N}\frac{1}{n!}$ is the N^{th} partial sum, then

$$\begin{aligned}e-s_N &= \sum_{n=N+1}^{\infty}\frac{1}{n!}=\frac{1}{(N+1)!}+\frac{1}{(N+2)!}++\frac{1}{(N+3)!}...\\ &= \frac{1}{(N+1)!}\left\{1+\frac{1}{N+2}+\frac{1}{(N+3)\,(N+2)}+...\right\}\\ &< \frac{1}{(N+1)!}\left\{1+\frac{1}{N+2}+\left(\frac{1}{N+2}\right)^{2}+...\right\}\\ &= \frac{1}{(N+1)!}\frac{1}{1-\frac{1}{N+2}}=\frac{1}{(N+1)!}\frac{N+2}{N+1},\end{aligned}$$

by Lemma 14. Now suppose that e is rational, say $e=\frac{p}{q}$ where $p,q\in\mathbb{N}$. Since $n!$ divides $q!$ for $n\leq q$ we conclude that

$$q!e-q!s_q=q!\frac{p}{q}-q!\sum_{n=0}^{q}\frac{1}{n!}=(q-1)!p-\sum_{n=0}^{q}\frac{q!}{n!}$$

is a positive integer satisfying

$$q!e-q!s_q<q!\frac{1}{(q+1)!}\frac{q+2}{q+1}=\frac{q+2}{(q+1)^{2}}<1,$$

a contradiction. Thus we have proved:

Theorem 19. *e is an* irrational *number lying strictly between* 2 *and* 3.

To prove the next familiar theorem on Euler's number e, it is convenient to introduce the *limit superior* and *limit inferior* of a real-valued sequence $\{s_n\}_{n=1}^{\infty}$.

DEFINITION 23. *Suppose that* $s = \{s_n\}_{n=1}^{\infty}$ *is a real-valued sequence and let* E^* *be the set of subsequential limits of* s. *Define*

$$\limsup_{n\to\infty} s_n \equiv \sup E^* \text{ and } \liminf_{n\to\infty} s_n \equiv \inf E^*,$$

called the limit superior *and* limit inferior *of* s *respectively.*

Since E^* is closed we have *either* $\limsup_{n\to\infty} s_n = \infty$ *or* $\limsup_{n\to\infty} s_n \in E^*$. In the latter case $\limsup_{n\to\infty} s_n$ is the largest subsequential limit of s. A similar comment applies to $\liminf_{n\to\infty} s_n$. Here are some easily verified properties of limit superior and limit inferior:

$$\begin{aligned} \liminf_{n\to\infty} s_n &\leq \limsup_{n\to\infty} s_n, \\ \lim_{n\to\infty} s_n &= L \text{ if and only if } \limsup_{n\to\infty} s_n = \liminf_{n\to\infty} s_n = L, \\ \limsup_{n\to\infty} t_n &\leq \limsup_{n\to\infty} s_n \text{ if } t_n \leq s_n \text{ for all sufficiently large } n, \\ \liminf_{n\to\infty} t_n &\geq \liminf_{n\to\infty} s_n \text{ if } t_n \geq s_n \text{ for all sufficiently large } n. \end{aligned} \tag{2.8}$$

THEOREM 20. $\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n = e$.

Proof: Let $s_n = \sum_{k=0}^{n} \frac{1}{k!}$ and $t_n = \left(1+\frac{1}{n}\right)^n$ for $n \geq 1$. By the binomial theorem

$$\begin{aligned} t_n &= \left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n} \frac{n!}{(n-k)!k!}\left(\frac{1}{n}\right)^k \\ &= 1 + \frac{n}{1!}\left(\frac{1}{n}\right) + \frac{n(n-1)}{2!}\left(\frac{1}{n}\right)^2 + \frac{n(n-1)(n-2)}{3!}\left(\frac{1}{n}\right)^3 + \ldots\left(\frac{1}{n}\right)^n \\ &= 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \ldots\frac{1}{n!}\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{n-1}{n}\right), \end{aligned}$$

and so $t_n \leq 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots\frac{1}{n!} = s_n$. Thus from the third line in (2.8) we have

$$\limsup_{n\to\infty} t_n \leq \limsup_{n\to\infty} s_n = e.$$

Conversely, fix $m > 1$. For $n > m$ we have

$$\begin{aligned} t_n &= 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \ldots\frac{1}{n!}\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{n-1}{n}\right) \\ &> 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \ldots\frac{1}{m!}\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{m-1}{n}\right). \end{aligned}$$

Now the limit as $n \to \infty$ of the last sum (remember that m is kept fixed) is

$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots\frac{1}{m!} = s_m.$$

Thus from the fourth line in (2.8) we have

$$\liminf_{n\to\infty} t_n \geq s_m$$

for all $m > 1$. Now take the limit as $m \to \infty$ to obtain $\liminf_{n\to\infty} t_n \geq e$. Altogether, using the first line in (2.8), we now have

$$e \leq \liminf_{n\to\infty} t_n \leq \limsup_{n\to\infty} t_n \leq e,$$

which implies that $\liminf_{n\to\infty} t_n = \limsup_{n\to\infty} t_n = e$. The second line in (2.8) now yields that $\lim_{n\to\infty} t_n = e$ as required.

3. Power series

There is a very special class of series that turn out to define complex-valued functions on balls in the complex plane. These are the so-called *power series* that have the form

$$\sum_{n=0}^{\infty} a_n z^n,$$

where $\{a_n\}_{n=0}^{\infty}$ is a sequence in $\mathbb{C}$ whose terms a_n are called *coefficients*, and where $z \in \mathbb{C}$ is called the *variable*. The first question of interest is: For what values of z in the complex plane does the series $\sum_{n=0}^{\infty} a_n z^n$ converge? The second question is: Of what use are these functions? The answer to the first question is initially surprising - namely the set of convergence E is either $\{0\}$, $\mathbb{C}$ or there is a ball $B(0,R)$ centered at the origin 0 with positive radius R such that

$$B(0,R) \subset E \subset \overline{B(0,R)}.$$

The answer to the second question is that these power series functions have many special properties, and moreover, *every* complex-valued function f defined on a ball $B(0,R)$ in $\mathbb{C}$ that has a *derivative* everywhere in $B(0,R)$ (i.e. $\lim_{w\to z} \frac{f(w)-f(z)}{w-z}$ exists for all $z \in B(0,R)$) turns out to be one of these power series functions! In other words

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad z \in B(0,R),$$

for some sequence of coefficients $\{a_n\}_{n=0}^{\infty}$. It turns out such f are infinitely differentiable and the coefficients are given by $a_n = \frac{f^{(n)}(0)}{n!}$. Many more magical properties of these so-called *analytic functions* are usually investigated in a course on complex analysis.

We content ourselves here with answering just the first question. This will require a new convergence test, the *root test*. We will also prove a close cousin, the *ratio test*.

THEOREM 21. *(Root Test) Let $\{a_n\}_{n=0}^{\infty}$ be a sequence of complex numbers and set*

$$L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}.$$

1. *If $L < 1$ then $\sum_{n=0}^{\infty} a_n$ converges,*
2. *If $L > 1$ then $\sum_{n=0}^{\infty} a_n$ diverges,*
3. *If $L = 1$ then there is no information.*

Proof: (1) Pick $L < R < 1$. Then there are only *finitely* many n satisfying $\sqrt[n]{|a_n|} \geq R$ (otherwise we would have $\limsup_{n\to\infty} \sqrt[n]{|a_n|} \geq R$), so there is N such that

$$\sqrt[n]{|a_n|} \leq R, \text{ i.e. } |a_n| \leq R^n, \qquad \text{for all } n \geq N.$$

Since $\sum_{n=0}^{\infty} R^n = \frac{1}{1-R}$ converges by Lemma 14, the comparison test Theorem 17 (1) shows that $\sum_{n=0}^{\infty} a_n$ converges.

(2) Since $1 < L$, there are *infinitely* many n satisfying $\sqrt[n]{|a_n|} \geq 1$ (otherwise we would have $\limsup_{n\to\infty} \sqrt[n]{|a_n|} \leq 1$), so there is N such that

$$\sqrt[n]{|a_n|} \geq 1, \text{ i.e. } |a_n| \geq 1, \qquad \text{for all } n \geq N.$$

Thus we cannot have $|a_n| \to 0$ as $n \to \infty$, and it follows that $\sum_{n=0}^{\infty} a_n$ diverges.

(3) The class of p-series shows that the root test gives no information on convergence when $L = 1$. Indeed, if $p > 0$, then

$$\limsup_{n\to\infty} \sqrt[n]{\frac{1}{n^p}} = \limsup_{n\to\infty} \left(\frac{1}{\sqrt[n]{n}}\right)^p = 1$$

by limit (3) at the beginning of the previous section. Yet for $p < 1$ the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ diverges and for $p > 1$ the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges.

COROLLARY 9. *Let $\{a_n\}_{n=0}^{\infty}$ be a sequence in $\mathbb{C}$ and set $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. Let $R = \frac{1}{L}$ (where $R = 0$ if $L = \infty$ and $R = \infty$ if $L = 0$). Then the set of convergence*

$$E = \left\{ z \in \mathbb{C} : \sum_{n=0}^{\infty} a_n z^n \text{ converges} \right\},$$

satisfies one of the following:

(1) *$E = \{0\}$ if $R = 0$,*
(2) *$E = \mathbb{C}$ if $R = \infty$,*
(3) *$B(0,R) \subset E \subset \overline{B(0,R)}$ if $0 < R < \infty$.*

The extended real number R is called the *radius of convergence* of the power series $\sum_{n=0}^{\infty} a_n z^n$.

Proof: Apply the root test to the series $\sum_{n=0}^{\infty} a_n z^n$ for $z \in \mathbb{C}$. We have

$$L = \limsup_{n\to\infty} \sqrt[n]{|a_n z^n|} = |z| \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|z|}{R}.$$

Thus if $z \in B(0,R)$, then $L < 1$ and the series $\sum_{n=0}^{\infty} a_n z^n$ converges, i.e. $z \in E$. If $z \notin \overline{B(0,R)}$, then $L > 1$ and the series $\sum_{n=0}^{\infty} a_n z^n$ diverges, i.e. $z \notin E$. This proves assertion (3), and the first two assertions are proved in similar fashion.

There is another test, the *ratio test*, that is often simpler to apply than the root test, but fails to have as wide a scope as the root test.

THEOREM 22. *(Ratio Test) Let $\{a_n\}_{n=0}^{\infty}$ be a sequence of complex numbers.*

(1) *If $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$ then $\sum_{n=0}^{\infty} a_n$ converges.*
(2) *If there is N such that $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ for all $n \geq N$, then $\sum_{n=0}^{\infty} a_n$ diverges.*

REMARK 11. *If $L = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$ exists, then $\sum_{n=0}^{\infty} a_n$ converges if $L < 1$, and $\sum_{n=0}^{\infty} a_n$ diverges if $L > 1$.*

Proof: (1) Pick $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < R < 1$. Then there are only *finitely* many n satisfying $\left|\frac{a_{n+1}}{a_n}\right| \geq R$ (otherwise we would have $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \geq R$), so there is N such that

$$\left|\frac{a_{n+1}}{a_n}\right| \leq R, \text{ i.e. } |a_{n+1}| \leq R\,|a_n|, \qquad \text{for all } n \geq N.$$

By induction we obtain

$$|a_{N+k}| \leq R\,|a_{N+k-1}| \leq R^2\,|a_{N+k-2}| \leq ... \leq R^k\,|a_N|\,, \qquad k \geq 0.$$

Now $\sum_{k=0}^{\infty} R^k\,|a_N| = \frac{|a_N|}{1-R}$ by Lemma 14, and so the comparison test Theorem 17 (1) shows that $\sum_{k=0}^{\infty} a_{N+k}$ converges, hence also $\sum_{n=0}^{\infty} a_n$.

(2) By induction we have

$$|a_{N+k}| \geq |a_{N+k-1}| \geq ... \geq |a_N|\,, \qquad k \geq 0.$$

Thus we cannot have $|a_n| \to 0$ as $n \to \infty$ and so $\sum_{n=0}^{\infty} a_n$ diverges.

PROBLEM 1. *What is the radius of convergence of the power series*

$$\sum_{n=0}^{\infty} \left(\begin{array}{c} 2n \\ n \end{array} \right) z^n = \sum_{n=0}^{\infty} \frac{(2n)!}{n!n!} z^n \quad ?$$

The root test is very hard to apply here without Stirling's formula $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$. *On the other hand the ratio test applies easily:*

$$\begin{aligned} \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| &= \lim_{n\to\infty} \left| \frac{\frac{(2n+2)!}{(n+1)!(n+1)!} z^{n+1}}{\frac{(2n)!}{n!n!} z^n} \right| \\ &= \lim_{n\to\infty} \left| \frac{(2n+2)\,(2n+1)}{(n+1)\,(n+1)} z \right| = 4\,|z|\,. \end{aligned}$$

By the remark following the ratio test, the power series converges if $|z| < \frac{1}{4}$ *and diverges if* $|z| > \frac{1}{4}$. *Thus the radius of convergence is* $\frac{1}{4}$.

PROBLEM 2. *What is the radius of convergence of the power series* $\sum_{n=0}^{\infty} \frac{z^n}{n!}$*? Since*

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \left| \frac{\frac{z^{n+1}}{(n+1)!}}{\frac{z^n}{n!}} \right| = \lim_{n\to\infty} \left| \frac{z}{n+1} \right| = 0,$$

we see that the radius of convergence is ∞. *This is the exponential function*

$$Exp\,(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \qquad z \in \mathbb{C}. \tag{3.1}$$

Finally, we note the sense in which the scope of the ratio test is not as wide as that of the root test.

PROPOSITION 11. *For any sequence* $\{a_n\}_{n=1}^{\infty}$ *of positive numbers we have*

$$\lim\sup_{n\to\infty} \sqrt[n]{a_n} \leq \lim\sup_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

Thus the root test gives convergence of the series $\sum_{n=0}^{\infty} a_n$ whenever the ratio test does.

Proof: Suppose $L \equiv \limsup_{n\to\infty} \frac{a_{n+1}}{a_n} < \infty$ and choose $L < R < \infty$. Then there is N such that

$$a_{n+1} \leq R a_n, \qquad n \geq N.$$

By induction we have

$$a_{N+k} \leq R a_{N+k-1} \leq R^2 a_{N+k-2} \leq ... \leq R^k a_N, \qquad k \geq 0,$$

and so with $n = N + k$,

$$\sqrt[n]{a_n} = (a_{N+k})^{\frac{1}{n}} \le \left(R^k \, |a_N|\right)^{\frac{1}{n}} = R^{\frac{k}{N+k}} a_N^{\frac{1}{n}} = R^{1-\frac{N}{N+k}} a_N^{\frac{1}{n}} = R^{1-\frac{N}{n}} a_N^{\frac{1}{n}}.$$

Now we take the limit superior as $n \to \infty$ to obtain

$$\limsup_{n\to\infty} \sqrt[n]{a_n} \le R \limsup_{n\to\infty} R^{-\frac{N}{n}} a_N^{\frac{1}{n}} = R \limsup_{n\to\infty} \left(\frac{a_N}{R^N}\right)^{\frac{1}{n}} = R$$

by limit (2) at the beginning of the previous subsection. Since $R > L$ was arbitrary we conclude that $\limsup_{n\to\infty} \sqrt[n]{a_n} \le L = \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}$ as required.

CHAPTER 5

Continuity and Differentiability

The notion of a continuous function $f : X \to Y$ makes sense when the function is defined from one metric space X to another Y. We will initially examine the connection between continuity and sequences, and after that between continuity and open sets. The notion of a differentiable function $f : X \to Y$ requires that X and Y be Euclidean spaces, usually the real or complex numbers. Central to all of this is the concept of *limit* of a function.

Definition 24. *Suppose that (X, d_X) and (Y, d_Y) are metric spaces. Let E be a subset of X and suppose that $f : E \to Y$ is a function from E to Y. Let $p \in X$ be a limit point of E and suppose that $q \in Y$. Then*

$$\lim_{x \to p} f(x) = q$$

if for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$d_Y(f(x), q) < \varepsilon \text{ whenever } x \in E \setminus \{p\} \text{ and } d_X(x, p) < \delta. \tag{0.2}$$

Note that the concept of a limit of f at a point p is only defined when p is a *limit point* of the set E on which f is defined. Do not confuse this notion with the definition of limit of a sequence $s = \{s_n\}_{n=1}^{\infty}$ in a metric space Y. In this latter definition, s is a function from the natural numbers $\mathbb{N}$ into the metric space Y, but the limit point p is replaced by the symbol ∞. Here is a characterization of limit of a function in terms of limits of sequences.

Theorem 23. *Suppose that (X, d_X) and (Y, d_Y) are metric spaces. Let E be a subset of X and suppose that $f : E \to Y$ is a function from E to Y. Let $p \in X$ be a limit point of E and suppose that $q \in Y$. Then* $\lim_{x \to p} f(x) = q$ if and only if

$$\lim_{n \to \infty} f(s_n) = q$$

for all sequences $\{s_n\}_{n=1}^{\infty}$ in $E \setminus \{p\}$ such that

$$\lim_{n \to \infty} s_n = p.$$

Proof: Suppose first that $\lim_{x \to p} f(x) = q$. Now assume that $\{s_n\}_{n=1}^{\infty}$ is a sequence in $E \setminus \{p\}$ such that $\lim_{n \to \infty} s_n = p$. Then given $\varepsilon > 0$ there is $\delta > 0$ such that (0.2) holds. Furthermore we can find N so large that $d_X(s_n, p) < \delta$ whenever $n \geq N$. Combining inequalities with the fact that $s_n \in E$ gives

$$d_Y(f(s_n), q) < \varepsilon \text{ whenever } n \geq N,$$

which proves $\lim_{n \to \infty} f(s_n) = q$.

Suppose next that $\lim_{x \to p} f(x) = q$ fails. The negation of Definition 24 is that **there exists** an $\varepsilon > 0$ such that **for every** $\delta > 0$ we have

$$d_Y(f(x), q) \geq \varepsilon \textbf{ for some } x \in E \setminus \{p\} \text{ with } d_X(x, p) < \delta. \tag{0.3}$$

So fix such an $\varepsilon > 0$ and for each $\delta = \frac{1}{n} > 0$ choose a point $s_n \in E \setminus \{p\}$ with $d_X(s_n, p) < \frac{1}{n}$. Then $\{s_n\}_{n=1}^{\infty}$ is a sequence in $E \setminus \{p\}$ such that the sequence $\{f(s_n)\}_{n=1}^{\infty}$ does not converge to q - indeed, $d_Y(f(s_n), q) \geq \varepsilon > 0$ for all $n \geq 1$.

As a corollary of the theorem we immediately obtain that limits are *unique* if they exist. In addition, if $Y = \mathbb{C}$ is the space of complex numbers, then limits behave as expected with regard to addition and multiplication.

Proposition 12. *Suppose that (X, d) is metric space. Let E be a subset of X and suppose that $f, g : E \to \mathbb{C}$ are complex-valued functions on E. Let $p \in X$ be a limit point of E and suppose that $A, B \in \mathbb{C}$ satisfy*

$$\lim_{x \to p} f(x) = A \text{ and } \lim_{x \to p} g(x) = B.$$

Then

(1) $\lim_{x \to p} \{f(x) + g(x)\} = A + B$.
(2) $\lim_{x \to p} f(x) g(x) = AB$.
(3) $\lim_{x \to p} \frac{f(x)}{g(x)} = \frac{A}{B}$ *provided* $B \neq 0$.

1. Continuous functions

A function $f : X \to Y$ from one metric space X to another Y is said to be continuous if it is continuous at each point p in X. We thus turn first to the definition of continuity at a point, which we give initially in a more general setting.

Definition 25. *Suppose that (X, d_X) and (Y, d_Y) are metric spaces. Let E be a subset of X and suppose that $f : E \to Y$ is a function from E to Y. Let $p \in E$. Then f is continuous at p if for every $\varepsilon > 0$ there is $\delta > 0$ such that*

$$d_Y(f(x), f(p)) < \varepsilon \text{ whenever } x \in E \text{ and } d_X(x, p) < \delta. \tag{1.1}$$

Note that (1.1) says

$$f(B(p, \delta) \cap E) \subset B(f(p), \varepsilon). \tag{1.2}$$

There are only two possibilities for the point $p \in E$; *either* p is a limit point of E *or* p is isolated in E (a point x in E is isolated in E if there is a deleted ball $B'(x, r)$ that has empty intersection with E). In the case that p is a limit point of E, then f is continuous at p if and only if $\lim_{x \to p} f(x)$ exists and the limit is $f(p)$, i.e.

$$\lim_{x \to p} f(x) = f(p). \tag{1.3}$$

On the other hand, if p is an isolated point of E, then f is automatically continuous at p since (1.1) holds for all $\varepsilon > 0$ with $\delta = r$ where $B'(x, r) \cap E = \emptyset$. From these remarks together with Theorem 23, we immediately obtain the following characterization of continuity in terms of sequences.

Theorem 24. *Suppose that X and Y are metric spaces. Let E be a subset of X and suppose that $f : E \to Y$ is a function from E to Y. Let $p \in E$. Then f is continuous at p* if and only if

$$\lim_{n \to \infty} f(s_n) = f(p)$$

for all sequences $\{s_n\}_{n=1}^{\infty}$ in $E \setminus \{p\}$ such that

$$\lim_{n \to \infty} s_n = p.$$

Remark 12. *The theorem remains true if we permit the sequences* $\{s_n\}_{n=1}^{\infty}$ *to lie in* E *rather than in* $E \setminus \{p\}$.

Before continuing any further, we point out that our definition of continuity of $f : E \to Y$ at a point $p \in E \subset X$ has absolutely *nothing* to do with the complement $X \setminus E$ of the set E in the ambient space X. Thus the definition of continuity at a point is intrinsic in the sense that it doesn't matter what ambient space X we choose to contain E, and in fact we can just restrict attention to the case $X = E$ is a metric space in its own right. Note that the definition of limit in Definition 24 is *not* intrinsic since the limit point p may not lie in the set E.

Definition 26. *A function* $f : X \to Y$ *is said to be* continuous on X *if* f *is continuous at each point* $p \in X$.

The previous theorem says that $f : X \to Y$ is continuous if and only if $\lim_{n\to\infty} f(s_n) = f(p)$ for all sequences $\{s_n\}_{n=1}^{\infty}$ in X such that $\lim_{n\to\infty} s_n = p$. There is an alternate characterization of continuity of $f : X \to Y$ in terms of open sets which is particularly useful in connection with compact sets and continuity of inverse functions.

Theorem 25. *Suppose that* $f : X \to Y$ *is a function from a metric space* X *to a metric space* Y. *Then* f *is continuous on* X *if and only if*

$$f^{-1}(G) \text{ is open in } X \text{ for every } G \text{ that is open in } Y. \tag{1.4}$$

Corollary 10. *Suppose that* $f : X \to Y$ *is a continuous function from a compact metric space* X *to a metric space* Y. *Then* $f(X)$ *is compact.*

Corollary 11. *Suppose that* $f : X \to Y$ *is a continuous function from a compact metric space* X *to a metric space* Y. *If* f *is both* one-to-one *and* onto, *then the inverse function* $f^{-1} : Y \to X$ *defined by*

$$f^{-1}(y) = x \text{ where } x \text{ is the unique point in } X \text{ satisfying } f(x) = y,$$

is a continuous map.

Proof (of Corollary 10): If $\{G_\alpha\}_{\alpha\in A}$ is an open cover of $f(X)$, then $\{f^{-1}(G_\alpha)\}_{\alpha\in A}$ is an open cover of X, hence has a finite subcover $\{f^{-1}(G_{\alpha_k})\}_{k=1}^{N}$. But then $\{G_{\alpha_k}\}_{k=1}^{N}$ is a finite subcover of $f(X)$ since

$$f(X) \subset f\left(\bigcup_{k=1}^{N} f^{-1}(G_{\alpha_k})\right) \subset \bigcup_{k=1}^{N} f\left(f^{-1}(G_{\alpha_k})\right) \subset \bigcup_{k=1}^{N} G_{\alpha_k}.$$

Note that it is *not* in general true that $f^{-1}(f(G)) \subset G$.

Proof (of Corollary 11): Let G be an open subset of X. We must show that $(f^{-1})^{-1}(G)$ is open in Y. Note that since f is one-to-one and onto, we have $(f^{-1})^{-1}(G) = f(G)$. Now $G^c = X \setminus G$ is closed in X, hence compact, and so Corollary 10 shows that $f(G^c)$ is compact, hence closed in Y, so $f(G^c)^c$ is open in Y. But again using that f is one-to-one and onto shows that $f(G) = f(G^c)^c$, and so we are done.

Remark 13. *Compactness is essential in this corollary since the map*

$$f : [0, 2\pi) \to \mathbb{T} \equiv \{z \in \mathbb{C} : |z| = 1\} \text{ defined by } f(\theta) = e^{i\theta} = (\cos\theta, \sin\theta),$$

takes $[0, 2\pi)$ *one-to-one and onto* $\mathbb{T}$*, yet the inverse map fails to be continuous at* $z = 1$*. Indeed, for points* z *on the circle just* below 1*,* $f^{-1}(z)$ *is close to* 2π*, while* $f^{-1}(1) = 0$*.*

Proof (of Theorem 25): Suppose first that f is continuous on X. We must show that (1.4) holds. So let G be an open subset of Y. We must now show that for every $p \in f^{-1}(G)$ there is $r > 0$ (depending on p) such that $B(p,r) \subset f^{-1}(G)$. Fix $p \in f^{-1}(G)$. Since G is open and $f(p) \in G$ we can pick $\varepsilon > 0$ such that $B(f(p), \varepsilon) \subset G$. But then by the continuity of f there is $\delta > 0$ such that (1.2) holds, i.e. $f(B(p,\delta)) \subset B(f(p), \varepsilon) \subset G$. It follows that

$$B(p,\delta) \subset f^{-1}(f(B(p,\delta))) \subset f^{-1}(G).$$

Conversely suppose that (1.4) holds. We must show that f is continuous at every $p \in X$. So fix $p \in X$. We must now show that for every $\varepsilon > 0$ there is $\delta > 0$ such that (1.2) holds, i.e. $f(B(p,\delta)) \subset B(f(p), \varepsilon)$. Fix $\varepsilon > 0$. Since $B(f(p), \varepsilon)$ is open, we have that $f^{-1}(B(f(p), \varepsilon))$ is open by (1.4). Since $p \in f^{-1}(B(f(p), \varepsilon))$ there is thus $\delta > 0$ such that $B(p,\delta) \subset f^{-1}(B(f(p), \varepsilon))$. It follows that

$$f(B(p,\delta)) \subset f\left(f^{-1}(B(f(p), \varepsilon))\right) \subset B(f(p), \varepsilon).$$

Before specializing to the case where Y is the space of real or complex numbers, we show that continuity is stable under composition of maps. Continuity on a metric space is easily handled with the help of Theorem 25.

Theorem 26. *Suppose that* X, Y, Z *are metric spaces. If* $f : X \to Y$ *and* $g : Y \to Z$ *are both continuous maps, then so is the composition* $h = g \circ f : X \to Z$ *defined by*

$$h(x) = g(f(x)), \qquad x \in X.$$

Proof: If G is open in Z, then

$$h^{-1}(G) = f^{-1}\left(g^{-1}(G)\right)$$

is open since g continuous implies $g^{-1}(G)$ is open by Theorem 25, and then f continuous implies $f^{-1}\left(g^{-1}(G)\right)$ is open by Theorem 25. Thus h is continuous by Theorem 25.

Continuity at a point is also easily handled using Definition 25. We leave the proof of the following theorem to the reader.

Theorem 27. *Suppose that* X, Y, Z *are metric spaces. If* $p \in E \subset X$ *and* $f : E \to Y$ *is continuous at* p *and* $g : f(E) \to Z$ *is continuous at* $f(p)$*, then the composition* $h = g \circ f : E \to Z$ *is continuous at* p*.*

1.1. Real and complex-valued continuous functions. Proposition 12 established limit properties for sums and products of complex-valued functions, and some definition chasing easily leads to the following analogous result for continuous maps.

Proposition 13. *If* f *and* g *are continuous complex-valued functions on a metric space* X*, then so are the functions* $f + g$ *and* fg*. If in addition* g *never vanishes, then* $\frac{f}{g}$ *is also continuous on* X*.*

Here is an extremely useful consequence of Corollary 10 when the target space Y is the real numbers.

THEOREM 28. *Suppose that X is a compact metric space and $f : X \to \mathbb{R}$ is continuous. Then there exist points $p, q \in X$ satisfying*

$$f(p) = \sup f(X) \text{ and } f(q) = \inf f(X).$$

REMARK 14. *Compactness of X is essential here as evidenced by the following example. If X is the open interval $(0,1)$ and $f : (0,1) \to (0,1)$ is the identity map defined by $f(x) = x$, then f is continuous and*

$$\begin{aligned} \sup f((0,1)) &= \sup(0,1) = 1, \\ \inf f((0,1)) &= \inf(0,1) = 0. \end{aligned}$$

However, there are no points $p, q \in (0,1)$ satisfying either $f(p) = 1$ or $f(q) = 0$.

Proof (of Theorem 28): Corollary 10 shows that $f(X)$ is compact. Lemmas 4 and 6 now show that $f(X)$ is a closed and bounded subset of $\mathbb{R}$. Finally, Theorem 5 shows that $\sup f(X)$ exists and that $\sup f(X) \in f(X)$, i.e. there is $p \in X$ such that $\sup f(X) = f(p)$. Similarly there is $q \in X$ satisfying $\inf f(X) = f(q)$.

Now consider a complex-valued function $f : X \to \mathbb{C}$ on a metric space X, and let $u : X \to \mathbb{R}$ and $v : X \to \mathbb{R}$ be the real and imaginary parts of f defined by

$$\begin{aligned} u(x) &= \operatorname{Re} f(x) \equiv \frac{f(x) + \overline{f(x)}}{2}, \\ v(x) &= \operatorname{Im} f(x) \equiv \frac{f(x) - \overline{f(x)}}{2i}, \end{aligned}$$

for $x \in X$. It is easy to see that f is continuous at a point $p \in X$ if and only if each of u and v is continuous at p. Indeed, the inequalities

$$\max\{|a|, |b|\} \leq \sqrt{|a|^2 + |b|^2} \leq |a| + |b|$$

show that if (1.1) holds for f (with $E = X$), i.e.

$$d_{\mathbb{C}}(f(x), f(p)) < \varepsilon \text{ whenever } d_X(x,p) < \delta,$$

then it also holds with f replaced by u or by v:

$$\begin{aligned} d_{\mathbb{R}}(u(x), u(p)) &= |u(x) - u(p)| \\ &\leq \sqrt{|u(x) - u(p)|^2 + |v(x) - v(p)|^2} \\ &= d_{\mathbb{C}}(f(x), f(p)) < \varepsilon \\ &\quad \text{whenever } d_X(x,p) < \delta. \end{aligned}$$

Similarly, if (1.1) holds for both u and v then it holds for f but with ε replaced by 2ε:

$$\begin{aligned} d_{\mathbb{C}}(f(x), f(p)) &= \sqrt{|u(x) - u(p)|^2 + |v(x) - v(p)|^2} \\ &\leq |u(x) - u(p)| + |v(x) - v(p)| \\ &= d_{\mathbb{R}}(u(x), u(p)) + d_{\mathbb{R}}(v(x), v(p)) < 2\varepsilon \\ &\quad \text{whenever } d_X(x,p) < \delta. \end{aligned}$$

The same considerations apply equally well to Euclidean space $\mathbb{R}^n$ (recall that $\mathbb{C} = \mathbb{R}^2$ as metric spaces) and we have the following theorem. Recall that the dot product of two vectors $\mathbf{z} = (z_1, ..., z_n)$ and $\mathbf{w} = (w_1, ..., w_n)$ in $\mathbb{R}^n$ is given by $\mathbf{z} \cdot \mathbf{w} = \sum_{k=1}^n z_k w_k$.

Theorem 29. *Let X be a metric space and suppose $\mathbf{f} : X \to \mathbb{R}^n$. Let $f_k(x)$ be the component functions defined by $\mathbf{f}(x) = (f_1(x), ..., f_n(x))$ for $1 \leq k \leq n$.*

1. *The vector-valued function $\mathbf{f} : X \to \mathbb{R}^n$ is continuous at a point $p \in X$* if and only if *each component function $f_k : X \to \mathbb{R}$ is continuous at p.*
2. *If both $\mathbf{f} : X \to \mathbb{R}^n$ and $\mathbf{g} : X \to \mathbb{R}^n$ are continuous at p then so are $\mathbf{f} + \mathbf{g} : X \to \mathbb{R}^n$ and $\mathbf{f} \cdot \mathbf{g} : X \to \mathbb{R}$.*

Here are some simple facts associated with the component functions on Euclidean space.

- For each $1 \leq j \leq n$, the component function $\mathbf{w} = (w_1, ..., w_n) \to w_j$ is continuous from $\mathbb{R}^n$ to $\mathbb{R}$.
- The length function $\mathbf{w} = (w_1, ..., w_n) \to |\mathbf{w}|$ is continuous from $\mathbb{R}^n$ to $[0, \infty)$; in fact we have the so-called reverse triangle inequality:
$$||\mathbf{z}| - |\mathbf{w}|| \leq |\mathbf{z} - \mathbf{w}|, \qquad \mathbf{z}, \mathbf{w} \in \mathbb{R}^n.$$
- Every monomial function $\mathbf{w} = (w_1, ..., w_n) \to w_1^{k_1} w_2^{k_2} ... w_n^{k_n}$ is continuous from $\mathbb{R}^n$ to $\mathbb{R}$.
- Every polynomial $P(\mathbf{w}) = \sum_{k_1 + ... k_n \leq N} a_{k_1,...k_n} w_1^{k_1} w_2^{k_2} ... w_n^{k_n}$ is continuous from $\mathbb{R}^n$ to $\mathbb{R}$.

1.2. Uniform continuity. A function $f : X \to Y$ that is continuous from a metric space X to another metric space Y satisfies Definition 25 at each point p in X, namely for every $p \in X$ and $\varepsilon > 0$ there is $\delta_p > 0$ (note the dependence on p) such that (1.1) holds with $E = X$:

$$d_Y(f(x), f(p)) < \varepsilon \text{ whenever } d_X(x, p) < \delta_p. \tag{1.5}$$

In general we cannot choose $\delta > 0$ to be independent of p. For example, the function $f(x) = \frac{1}{x}$ is continuous on the open interval $(0, 1)$, but if we want

$$\varepsilon > d_Y(f(x), f(p)) = \left|\frac{1}{x} - \frac{1}{p}\right| \text{ whenever } |p - x| < \delta,$$

we cannot take $p = \delta$ since then x could be arbitrarily close to 0, and so $\frac{1}{x}$ could be arbitrarily large. In this example, $X = (0, 1)$ is not compact and this turns out to be the reason we cannot choose $\delta > 0$ to be independent of p. The surprising property that continuous functions f on a *compact* metric space X have is that we *can* indeed choose $\delta > 0$ to be independent of p in (1.5). We first give a name to this surprising property; we call it *uniform continuity* on X.

Definition 27. *Suppose that $f : X \to Y$ maps a metric space X into a metric space Y. We say that f is* uniformly continuous *on X if for every $\varepsilon > 0$ there is $\delta > 0$ such that*

$$d_Y(f(x), f(p)) < \varepsilon \textit{ whenever } d_X(x, p) < \delta.$$

The next theorem plays a crucial role in the theory of integration and its application to existence and uniqueness of solutions to differential equations.

Theorem 30. *Suppose that $f : X \to Y$ is a continuous map from a* compact *metric space X into a metric space Y. Then f is uniformly continuous on X.*

Proof: Suppose $\varepsilon > 0$. Since f is continous on X, (1.2) shows that for each point $p \in X$, there is $\delta_p > 0$ such that

$$f\left(B\left(p, \delta_p\right)\right) \subset B\left(f\left(p\right), \frac{\varepsilon}{2}\right). \tag{1.6}$$

Since X is compact, the open cover $\left\{B\left(p, \frac{\delta_p}{2}\right)\right\}_{p \in X}$ has a finite subcover $\left\{B\left(p_k, \frac{\delta_{p_k}}{2}\right)\right\}_{k=1}^N$. Now define

$$\delta = \min\left\{\frac{\delta_{p_k}}{2}\right\}_{k=1}^N.$$

Since the minimum is taken over *finitely* many positive numbers (thanks to the *finite* subcover, which in turn owes its existence to the *compactness* of X), we have $\delta > 0$.

Now suppose that $x, p \in X$ satisfy $d_X\left(x, p\right) < \delta$. We will show that

$$d_Y\left(f\left(x\right), f\left(p\right)\right) < \varepsilon.$$

Choose k so that $p \in B\left(p_k, \frac{\delta_{p_k}}{2}\right)$. Then we have using the triangle inequality in X that

$$d_X\left(x, p_k\right) \leq d_X\left(x, p\right) + d_X\left(p, p_k\right) < \delta + \frac{\delta_{p_k}}{2} \leq \frac{\delta_{p_k}}{2} + \frac{\delta_{p_k}}{2} = \delta_{p_k},$$

so that both p and x lie in the ball $B\left(p_k, \delta_{p_k}\right)$. It follows from (1.6) that both $f\left(p\right)$ and $f\left(x\right)$ lie in

$$f\left(B\left(p_k, \delta_{p_k}\right)\right) \subset B\left(f\left(p_k\right), \frac{\varepsilon}{2}\right).$$

Finally an application of the triangle inequality in Y shows that

$$d_Y\left(f\left(x\right), f\left(p\right)\right) \leq d_Y\left(f\left(x\right), f\left(p_k\right)\right) + d_Y\left(f\left(p_k\right), f\left(p\right)\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

1.3. Connectedness.

DEFINITION 28. *A metric space X is said to be* connected *if it is* not *possible to write $X = E \dot{\cup} F$ where E and F are disjoint nonempty open subsets of X. A subset Y of a metric space X is connected if it is connected when considered as a metric space in its own right. A set that is not connected is said to be* disconnected.

Equivalently, X is disconnected if it has a nonempty proper *clopen* subset (a clopen subset of X is one that is simultaneously open and closed in X).

LEMMA 16. *A subset Y of X is disconnected if and only if there are nonempty subsets E and F of X with $Y = E \dot{\cup} F$ and*

$$\overline{E} \cap F = \emptyset \text{ and } E \cap \overline{F} = \emptyset, \tag{1.7}$$

where the closures refer to the ambient metric space X.

Proof: Theorem 4 shows that E is an open subset of the metric space Y if and only if $E \cap \overline{F} = \emptyset$. Similarly, F is open in Y if and only if $\overline{E} \cap F = \emptyset$. Finally, E is clopen in Y if and only if both E and $F = Y \setminus E$ are open in Y.

The connected subsets of the real line are especially simple - they are precisely the intervals

$$[a, b], (a, b), [a, b), (a, b]$$

lying in $\mathbb{R}$ with $-\infty \le a \le b \le \infty$ (we do not consider any case where a or b is $\pm\infty$ and lies next to either [or]).

THEOREM 31. *The connected subsets of the real numbers* $\mathbb{R}$ *are precisely the intervals.*

Proof: Consider first a nonempty connected subset Y of $\mathbb{R}$. If $a, b \in Y$, and $a < c < b$, then we must also have $c \in Y$ since otherwise $Y \cap (-\infty, c)$ is clopen in Y. Thus the set Y has the *intermediate value property* ($a, b \in Y$ and $a < c < b$ implies $c \in Y$), and it is now easy to see using the Least Upper Bound Property of $\mathbb{R}$, that Y is an interval. Conversely, if Y is a disconnected subset of $\mathbb{R}$, then Y has a nonempty proper clopen subset E. We can then find two points $a, b \in Y$ with $a \in E$ and $b \in F \equiv Y \setminus E$ and (without loss of generality) $a < b$. Set

$$c \equiv \sup (E \cap [a, b]) .$$

By Theorem 5 we have $c \in \overline{E}$, and so $c \notin F$ by (1.7). If also $c \notin E$, then Y fails the intermediate value property and so cannot be an interval. On the other hand, if $c \in E$ then $c \notin \overline{F}$ (the *closure* of F), and so there is $d \in (c, b) \setminus F$. But then $d \notin E$ since $d > c$ and so lies in $(a, b) \setminus Y$, which again shows that Y fails the intermediate value property and so cannot be an interval.

Connected sets behave the same way as compact sets under pushforward by a continuous map.

THEOREM 32. *Suppose* $f : X \to Y$ *is a continuous map from a metric space* X *to another metric space* Y*, and suppose that* A *is a subset of* X*. If* A *is connected, then* $f(A)$ *is connected.*

Proof: We may suppose that $A = X$ and $f(A) = Y$. If Y is disconnected, there are disjoint nonempty open subsets E and F with $Y = E \dot{\cup} F$. But then $X = f^{-1}(E) \dot{\cup} f^{-1}(F)$ where both $f^{-1}(E)$ and $f^{-1}(F)$ are open in X by Theorem 25. This shows that X is disconnected as well, and completes the proof of the (contrapositive of the) theorem.

COROLLARY 12. *If* $f : \mathbb{R} \to \mathbb{R}$ *is continuous, then* f *takes intervals to intervals, and in particular,* f *takes closed bounded intervals to closed bounded intervals.*

Note that this corollary yields two familiar theorems from first year calculus, the Intermediate Value Theorem (real continuous functions on an interval attain their intermediate values) and the Extreme Value Theorem (real continuous functions on a closed bounded interval attain their extreme values).

Proof: Apply Theorems 32, 11 and 10.

Finally we have the following simple description of open subsets of the real numbers.

PROPOSITION 14. *Every open subset* G *of the real numbers* $\mathbb{R}$ *can be uniquely written as an at most countable pairwise disjoint union of open intervals* $\{I_n\}_{n\ge1}$*:*

$$G = \dot{\bigcup_{n\ge1}} I_n .$$

Proof: For $x \in G$ let

$$I_x = \bigcup \{\text{all open intervals containing } x \text{ that are contained in } G\}.$$

It is easy to see that I_x is an open interval and that if $x, y \in G$ then

$$\text{either } I_x = I_y \text{ or } I_x \cap I_y = \emptyset.$$

This shows that G is a union $\bigcup_{\alpha \in A} I_\alpha$ of pairwise disjoint open intervals. To see that this union is at most countable, simply use (2) of Proposition 3 to pick a rational number r_α in each I_α. The uniqueness is left as an exercise for the reader.

2. Differentiable functions

We can define the derivative of a real-valued function f at a point p provided f is defined on an *interval* I containing p. We give the definition when I is a closed interval, the remaining cases being similar.

DEFINITION 29. *Suppose $f : [a,b] \to \mathbb{R}$ and that $p \in [a,b]$. Then p is a limit point of $E \equiv [a,b] \setminus \{p\}$ and the function $Q(x) = \frac{f(x)-f(p)}{x-p}$ of Difference Quotients is defined on E. We say that f is* differentiable at x *if there is $q \in \mathbb{R}$ such that*

$$\lim_{x \to p} Q(x) = q$$

in accordance with Definition 24. In this case we say that q is the derivative of f at p and we write

$$f'(p) \equiv q = \lim_{x \to p} Q(x) = \lim_{x \to p} \frac{f(x) - f(p)}{x - p}. \tag{2.1}$$

In the case $p = a$, we say that $f'(a)$ defined as above is a *right hand* derivative of f at a, while if $p = b$, we that $f'(b)$ is a *left hand* derivative of f at b. We can of course define left and right hand derivatives of f at $p \in (a,b)$ by restricting the domain of f to $[a,p]$ and $[p,b]$ respectively. If f is differentiable at every point in a subset E of $[a,b]$, then we say that f is *differentiable* on E.

REMARK 15. *The Difference Quotient $\frac{f(x)-f(p)}{x-p}$ is the slope of the line segment joining the points $(p, f(p))$ and $(x, f(x))$ on the graph of f. Thus if $f'(p)$ exists, it is the limiting value of the slopes of the line segements $\overrightarrow{(p, f(p))(x, f(x))}$ as $x \to p$, and so we* define *the line L through the point $(p, f(p))$ having this limiting slope $f'(p)$ to be the* tangent line *to the graph of f at the point $(p, f(p))$. The equation of the tangent line L is*

$$y = f(p) + f'(p)(x - p), \qquad x \in \mathbb{R}. \tag{2.2}$$

LEMMA 17. *Suppose $f : [a,b] \to \mathbb{R}$ and that $p \in [a,b]$. If f is differentiable at p, then f is continuous at p.*

Proof: We have

$$\begin{aligned}
\lim_{x \to p} (f(x) - f(p)) &= \lim_{x \to p} \left(\frac{f(x) - f(p)}{x - p} \right) (x - p) \\
&= \left(\lim_{x \to p} \frac{f(x) - f(p)}{x - p} \right) \left(\lim_{x \to p} (x - p) \right) \\
&= f'(p) \cdot 0 = 0,
\end{aligned}$$

which implies $\lim_{x \to p} f(x) = f(p)$. Thus f is continuous at p by (1.3).

Now we investigate the calculus of derivatives. First we have the derivative calculus of the field operations. To state the formulas we revert to the more common notation of using x in place of p as the point at which we compute derivatives.

PROPOSITION 15. *Suppose that $f, g : (a, b) \to \mathbb{R}$ are functions differentiable at a point $x \in (a, b)$, and suppose that $c \in \mathbb{R}$ represents the constant function. Then we have*

(1) $(f + g)'(x) = f'(x) + g'(x)$,
(2) $(cf)'(x) = cf'(x)$,
(3) $(fg)'(x) = f'(x) g(x) + f(x) g'(x)$,
(4) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$ *provided* $g(x) \neq 0$.

Proof: For example, to prove (3) we use (2.1) and the corresponding properties of limits to obtain

$$\begin{aligned}
(fg)'(x) &= \lim_{y \to x} \frac{(fg)(y) - (fg)(x)}{y - x} \\
&= \lim_{y \to x} \left\{ \frac{f(y) g(y) - f(x) g(y)}{y - x} + \frac{f(x) g(y) - f(x) g(x)}{y - x} \right\} \\
&= \lim_{y \to x} \frac{f(y) - f(x)}{y - x} \lim_{y \to x} g(y) + f(x) \lim_{y \to x} \frac{g(y) - g(x)}{y - x} \\
&= f'(x) g(x) + f(x) g'(x).
\end{aligned}$$

The other formulas are proved similarly.

Second we have the calculus of composition of functions, the so-called "chain rule". This is most easily proved using an equivalent formulation of differentiability due to Landau. We begin by rewriting (2.1) in the alternate form

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}.$$

Then we rewrite this latter expression using Landau's "little oh" notation as

$$f(x + h) = f(x) + f'(x) h + o(h), \tag{2.3}$$

where $o(h)$ denotes a function of h satisfying $\frac{o(h)}{h} \to 0$ as $h \to 0$.

PROPOSITION 16. *Suppose that f is differentiable at x and that g is differentiable at $y = f(x)$. Then*

$$(g \circ f)'(x) = g'(y) f'(x) = g'(f(x)) f'(x).$$

Proof: We use the Landau formulation (2.3) of derivative and the corresponding properties of limits as follows. Write

$$\begin{aligned}
f(x + h_1) &= f(x) + f'(x) h_1 + o_1(h_1), \\
g(y + h_2) &= g(y) + g'(y) h_2 + o_2(h_2),
\end{aligned}$$

and then with

$$h_2 = f(x + h_1) - f(x) = f'(x) h_1 + o_1(h_1),$$

we have,

$$\begin{aligned} (g \circ f)(x+h_1) &= g(f(x+h_1)) \\ &= g(f(x) + f'(x)h_1 + o_1(h_1)) \\ &= g(y+h_2) \\ &= g(y) + g'(y)h_2 + o_2(h_2) \\ &= (g\circ f)(x) + g'(y)\{f'(x)h_1 + o_1(h_1)\} + o_2(h_2) \\ &= (g\circ f)(x) + g'(y)f'(x)h_1 + o_3(h_1) \end{aligned}$$

where using $\lim_{h_1\to 0} h_2 = 0$, we conclude that as $h_1 \to 0$,

$$\frac{o_3(h_1)}{h_1} \equiv g'(y)\frac{o_1(h_1)}{h_1} + \frac{o_2(h_2)}{h_2}\frac{f'(x)h_1 + o_1(h_1)}{h_1} \to g'(y)\cdot 0 + 0 \cdot f'(x) = 0.$$

EXAMPLE 8. *There is a function $f : \mathbb{R} \to \mathbb{R}$ whose derivative $f' : \mathbb{R} \to \mathbb{R}$ exists everywhere on the real line, but the derivative function f' is* not *itself differentiable at 0, not even continuous at 0. For example*

$$f(x) = \begin{cases} x^2 \sin\frac{1}{x} & \text{if} \quad x \neq 0 \\ 0 & \text{if} \quad x = 0 \end{cases}$$

has these properties. Indeed,

$$f'(x) = \begin{cases} 2x\sin\frac{1}{x} - \cos\frac{1}{x} & \text{if} \quad x \neq 0 \\ 0 & \text{if} \quad x = 0 \end{cases}$$

fails to be continuous at the origin.

PROPOSITION 17. *Suppose that $f : [a,b] \to \mathbb{R}$ is continuous and strictly increasing. Let $x \in (a,b)$ and set $y = f(x)$. If f is differentiable at x and $f'(x) \neq 0$, then f^{-1} is differentiable at y and*

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}.$$

Proof: We first note that by Corollary 12, $f : [a,b] \to [f(a), f(b)]$ is continuous, one-to-one and onto. Thus Corollary 11 shows that f^{-1} is continuous. Then with

$$h = f^{-1}(y+k) - f^{-1}(y) = f^{-1}(y+k) - x,$$

we have

$$f(x+h) = f\left(f^{-1}(y+k)\right) = y + k,$$

and so

$$\frac{f^{-1}(y+k) - f^{-1}(y)}{k} = \frac{h}{f(x+h) - f(x)} \to \frac{1}{f'(x)}$$

as $k \to 0$ since $f'(x) \neq 0$ and

$$\lim_{k\to 0} h = \lim_{k \to 0}\left(f^{-1}(y+k) - f^{-1}(y)\right) = 0$$

by the continuity of f^{-1} at y.

2.1. Mean value theorems. We will present four mean value theorems in order of increasing generality. They all depend on the following theorem of Fermat. If $f : X \to \mathbb{R}$ where X is any metric space, we say that f has a *relative maximum* at a point p in X if there is $\delta > 0$ such that

$$f(p) \geq f(x) \text{ for all } x \in B(p, \delta).$$

A relative minimum is defined similarly.

THEOREM 33. *Suppose $f : [a,b] \to \mathbb{R}$ and $p \in (a,b)$. If f has either a relative maximum or a relative minimum at p, and if f is differentiable at p, then*

$$f'(p) = 0.$$

Proof: Suppose f has a relative maximum at p. Then there is $\delta > 0$ such that $f(x) - f(p) \leq 0$ for $x \in (p - \delta, p + \delta)$. It follows that

$$\begin{aligned} \frac{f(x) - f(p)}{x - p} &\leq 0, \qquad \text{for } x \in (p, p+\delta), \\ \frac{f(x) - f(p)}{x - p} &\geq 0, \qquad \text{for } x \in (p-\delta, p). \end{aligned}$$

If we take a sequence $\{x_n\}_{n=1}^{\infty}$ in $(p, p+\delta)$ converging to p, we see that

$$f'(p) = \lim_{n\to\infty} \frac{f(x_n) - f(p)}{x_n - p} \leq 0,$$

and if we take a sequence $\{x_n\}_{n=1}^{\infty}$ in $(p - \delta, p)$ converging to p, we see that

$$f'(p) = \lim_{n\to\infty} \frac{f(x_n) - f(p)}{x_n - p} \geq 0.$$

Combining these inequalities proves that $f'(p) = 0$. The proof is similar if f has a relative minimum at p.

THEOREM 34. ***(First Mean Value)*** *Suppose that $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b). If $f(a) = f(b) = 0$, then there is $c \in (a,b)$ such that*

$$f'(c) = 0.$$

Proof: If $f \equiv 0$ then any $c \in (a,b)$ works. Otherwise we may suppose without loss of generality that $f(x) > 0$ for some x. Then by Theorem 28 there is $c \in [a,b]$ such that

$$\sup f([a,b]) = f(c).$$

Since $f(c) \geq f(x) > 0$ we must have $c \in (a,b)$, and so f has a relative maximum at c. Theorem 33 now implies $f'(c) = 0$.

THEOREM 35. ***(Second Mean Value)*** *Suppose that $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b). Then there is $c \in (a,b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof: Define $g : [a,b] \to \mathbb{R}$ by

$$g(x) = f(x) - \left\{ f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \right\}, \qquad a \leq x \leq b,$$

so that $g(x)$ is the signed vertical distance from the graph of f at x to the graph of the line joining $(a, f(a))$ to $(b, f(b))$ at x. Then g satisfies the hypotheses of Theorem 34 and so there is a point $c \in (a, b)$ satisfying

$$0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}.$$

Note that the conclusion of the second mean value theorem can be rewritten as

$$f(b) = f(a) + f'(c)(b - a). \tag{2.4}$$

THEOREM 36. ***(Third Mean Value)*** *Suppose that $f, g : [a, b] \to \mathbb{R}$ are each continuous on $[a, b]$ and differentiable on (a, b). Then there is $c \in (a, b)$ such that*

$$[g(b) - g(a)] f'(c) = [f(b) - f(a)] g'(c).$$

Proof: Define $h : [a, b] \to \mathbb{R}$ by

$$h(x) = [g(b) - g(a)] f(x) - [f(b) - f(a)] g(x), \qquad a \le x \le b.$$

Then h satisfies the hypotheses of Theorem 35 and a small calculation shows that $h(a) = h(b)$. So there is a point $c \in (a, b)$ satisfying

$$0 = \frac{h(b) - h(a)}{b - a} = h'(c) = [g(b) - g(a)] f'(c) - [f(b) - f(a)] g'(c).$$

DEFINITION 30. *If $f : [a, b] \to \mathbb{R}$ is differentiable on $[a, b]$, and if $f' : [a, b] \to \mathbb{R}$ is differentiable on a subset E of $[a, b]$, then we define $f'' = (f')'$ on E, and call f'' the second derivative of f on E. More generally, for $n \ge 2$ we define $f^{(n)} = \left(f^{(n-1)}\right)'$ on E if $f^{(n-1)}$ is defined on an interval containing E.*

The form (2.4) can be generalized to higher order derivatives.

THEOREM 37. ***(Fourth Mean Value)*** *Suppose that $f : [a, b] \to \mathbb{R}$ is $n - 1$ times continuously differentiable on $[a, b]$, i.e. $f, f', ..., f^{(n-1)}$ are each defined and continuous on $[a, b]$, and suppose that $f^{(n-1)}$ is differentiable on (a, b), i.e. $f^{(n)}$ exists on (a, b). Then there is $c \in (a, b)$ such that*

$$\begin{aligned} f(b) &= f(a) + f'(a)(b - a) + f''(a) \frac{(b - a)^2}{2!} + ... + f^{(n-1)}(a) \frac{(b - a)^{n-1}}{(n - 1)!} \\ &\quad + f^{(n)}(c) \frac{(b - a)^n}{n!} \\ &= \sum_{k=0}^{n-1} f^{(k)}(a) \frac{(b - a)^k}{k!} + f^{(n)}(c) \frac{(b - a)^n}{n!}. \end{aligned}$$

Proof: Define $g : [a, b] \to \mathbb{R}$ by

$$g(x) = f(x) - \sum_{k=0}^{n-1} f^{(k)}(a) \frac{(x - a)^k}{k!} + M(x - a)^n, \qquad a \le x \le b,$$

and where M is the number uniquely defined by requiring $g(b) = 0$, i.e.

$$M(b - a)^n = \sum_{k=0}^{n-1} f^{(k)}(a) \frac{(b - a)^k}{k!} - f(b).$$

Calculations show that

$$
\begin{aligned}
(2.5)\qquad g'(x) &= f'(x) - \sum_{k=1}^{n-1} f^{(k)}(a) \frac{(x-a)^{k-1}}{(k-1)!} + nM(x-a)^{n-1}, \\
g''(x) &= f''(x) - \sum_{k=2}^{n-1} f^{(k)}(a) \frac{(x-a)^{k-2}}{(k-2)!} + n(n-1)M(x-a)^{n-2}, \\
&\vdots \\
g^{(n-1)}(x) &= f^{(n-1)}(x) - f^{(n-1)}(a) + n(n-1)\ldots(3)(2)(1)M(x-a), \\
g^{(n)}(x) &= f^{(n)}(x) - 0 + n!M.
\end{aligned}
$$

Now the conclusion of the theorem is that

$$f^{(n)}(c)\frac{(b-a)^n}{n!} = f(b) - \sum_{k=0}^{n-1} f^{(k)}(a)\frac{(b-a)^k}{k!} = -M(b-a)^n,$$

i.e. $f^{(n)}(c) + n!M = 0$. Thus using the last line in (2.5) we see that we must show $g^{(n)}(c) = 0$ for some $c \in (a,b)$.

Now the ℓ^{th} line of (2.5) shows that

$$(2.6)\qquad g^{(\ell)}(a) = f^{(\ell)}(a) - f^{(\ell)}(a) + 0 = 0, \qquad 0 \le \ell \le n-1.$$

Since $g(a) = g(b) = 0$, the first mean value theorem shows that there is $c_1 \in (a,b)$ satisfying

$$g'(c_1) = 0.$$

Using (2.6) we see that $g'(a) = g'(c_1) = 0$, and so the first mean value theorem shows that there is $c_2 \in (a, c_1)$ satisfying

$$g''(c_2) = 0.$$

Continuing in this way we obtain $c_\ell \in (a, c_{\ell-1})$ satisfying

$$g^{(\ell)}(c_\ell) = 0,$$

for each $1 \le \ell \le n$. The number $c = c_n \in (a,b)$ satisfies $g^{(n)}(c) = 0$ and this completes the proof of the fourth mean value theorem.

REMARK 16. *The first three mean value theorems can each be interpreted as saying that there is a point on a curve whose tangent is parallel to the line segment joining the endpoints of the curve. For example, in the second theorem, $f'(c)$ is the slope of the tangent line to the graph of f at $(c, f(c))$, while $\frac{f(b)-f(a)}{b-a}$ is the slope of the line joining the endpoints $(a, f(a))$ and $(b, f(b))$ of the graph. In the third theorem, $\frac{f'(c)}{g'(c)}$ is the slope of the parametric curve $x \to (f(x), g(x))$ at the point $(f(c), g(c))$, while $\frac{f(b)-f(a)}{g(b)-g(a)}$ is the slope of the line joining the endpoints $(f(a), g(a))$ and $(f(b), g(b))$. On the other hand, the second and fourth theorems can each be interpreted as saying that a function can be approximated by a polynomial.*

2.2. Some consequences of the mean value theorems.

THEOREM 38. *(monotone functions) Suppose* $f:(a,b)\to\mathbb{R}$ *is differentiable.*

(1) *If* $f'(x)=0$ *for all* $x\in(a,b)$*, then* f *is constant on* (a,b)*.*
(2) *If* $f'(x)\geq 0$ *(respectively* $f'(x)\leq 0$*) for all* $x\in(a,b)$*, then* f *is monotonically increasing (respectively decreasing) on* (a,b)*.*

Proof: Apply (2.4) of the second mean value theorem to the interval $[\alpha,\beta]$ for any $a<\alpha<\beta<b$ to obtain

$$f(\beta)=f(\alpha)+f'(c)(\beta-\alpha),$$

for some $c\in(\alpha,\beta)$.

(1) If $f'(c)=0$ for all $c\in(\alpha,\beta)$, then $f(\beta)=f(\alpha)$ for all $a<\alpha<\beta<b$.

(2) If $f'(c)\geq 0$ for all $c\in(\alpha,\beta)$, then $f(\beta)\geq f(\alpha)$ for all $a<\alpha<\beta<b$. If $f'(c)\leq 0$ for all $c\in(\alpha,\beta)$, then $f(\beta)\leq f(\alpha)$ for all $a<\alpha<\beta<b$.

Recall from Corollary 12 that continuous functions have the Intermediate Value Property. The next theorem shows that derivatives also have the Intermediate Value Property, despite the fact that they need not be continuous functions - see Example 8. This is often referred to as a continuity property of derivatives.

THEOREM 39. *(continuity of derivatives) Suppose* $f:[a,b]\to\mathbb{R}$ *is differentiable. If* $f'(a)<\lambda<f'(b)$*, then there is* $c\in(a,b)$ *such that* $f'(c)=\lambda$*.*

Proof: We effectively reduce matters to the case $\lambda=0$ by considering $g:[a,b]\to\mathbb{R}$ defined by

$$g(x)=f(x)-\lambda x,\qquad x\in[a,b].$$

By Theorem 28 there is a point $p\in[a,b]$ such that

$$\inf g([a,b])=g(p).$$

We claim that $p\in(a,b)$, i.e. that p cannot be either of the endpoints a or b. Indeed,

$$g'(x)=f'(x)-\lambda \tag{2.7}$$

and so

$$\begin{aligned} g'(a) &= f'(a)-\lambda<0,\\ g'(b) &= f'(b)-\lambda>0. \end{aligned}$$

Since $0>g'(a)=\lim_{x\to a}\frac{g(x)-g(a)}{x-a}$, there is some $x_1\in(a,b)$ such that

$$g(x_1)-g(a)<0,$$

and this shows that $p\neq a$. Since $0<g'(b)=\lim_{x\to b}\frac{g(x)-g(b)}{x-b}$, there is some $x_2\in(a,b)$ such that

$$g(x_2)-g(b)<0,$$

and this shows that $p\neq b$. Thus g has a *relative* minimum at p and by Theorem 33 we conclude that $g'(p)=0$. Hence $f'(p)=\lambda$ by (2.7).

THEOREM 40. *(l'Hôspital's rule) Suppose* $f,g:(a,b)\to\mathbb{R}$ *are each differentiable, and that* $g'(x)\neq 0$ *for all* $a<x<b$*. If* $\lim_{x\to a}f(x)=\lim_{x\to a}g(x)=0$ *and*

$$\lim_{x\to a}\frac{f'(x)}{g'(x)}=L,$$

then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = L.$$

Proof: Given $\varepsilon > 0$ there is $\delta > 0$ such that

$$\left| \frac{f'(x)}{g'(x)} - L \right| < \varepsilon \text{ for all } a < x < a + \delta.$$

Now for $a < \alpha < \beta < a + \delta$, the third mean value theorem gives a point $c \in (\alpha, \beta)$ such that

$$[g(\beta) - g(\alpha)] f'(c) = [f(\beta) - f(\alpha)] g'(c),$$

and since $g'(c) \neq 0$ we can write

$$\frac{f'(c)}{g'(c)} = \frac{f(\beta) - f(\alpha)}{g(\beta) - g(\alpha)}.$$

Thus we have

$$\left| \frac{f(\beta) - f(\alpha)}{g(\beta) - g(\alpha)} - L \right| < \varepsilon \text{ for all } a < \alpha < \beta < a + \delta.$$

Now let $\alpha \to a$ and use $\lim_{\alpha \to a} f(\alpha) = \lim_{\alpha \to a} g(\alpha) = 0$ to get

$$\left| \frac{f(\beta)}{g(\beta)} - L \right| \leq \varepsilon \text{ for all } a < \beta < a + \delta.$$

This completes the proof that $\lim_{x \to a} \frac{f(x)}{g(x)} = L$.

Part 2

Integration

In the second part of these notes we consider the problem of describing the inverse operation to that of differentiation, commonly called *integration*. There are four widely recognized theories of integration:

- Riemann integration - the workhorse of integration theory that provides us with the most basic form of the fundamental theorem of calculus;
- Riemann-Stieltjes integration - that extends the idea of integrating the infinitesmal dx to that of the more general infinitesmal $d\alpha(x)$ for an increasing function α.
- Lebesgue integration - that overcomes a shortcoming of the Riemann theory by permitting a robust theory of limits of functions, all at the expense of a complicated theory of 'measure' of a set.
- Henstock-Kurtzweil integration - that includes the Riemann and Lebesgue theories and has the advantages that it is quite similar in spirit to the intuitive Riemann theory, and avoids much of the complication of measurability of sets in the Lebesgue theory. However, it has the drawback of limited scope for generalization.

In Chapter 6 we follow Rudin [**3**] and use uniform continuity to develop the standard theory of the Riemann and Riemann-Stieltjes integrals. A short detour is taken to introduce the more powerful Henstock-Kurtzweil integral, and we use compactness to prove its uniqueness and extension properties.

In Chapter 7 we prove the familiar theorems on uniform convergence of functions and apply this to prove that the metric space $C_{\mathbb{R}}(X)$ of real-valued continuous functions on a compact metric space X is complete. We then use integration theory and the Contraction Lemma from Chapter 4 to produce an elegant proof of existence and uniqueness of solutions to certain initial value problems for differential equations. We also construct a space-filling curve and the von Koch snowflake.

Chapter 8 draws on Stein and Shakarchi [**5**] to provide a rapid introduction to the theory of the Lebesgue integral.

CHAPTER 6

Riemann and Riemann-Stieltjes integration

Let $f : [0,1] \to \mathbb{R}$ be a bounded function on the closed unit interval $[0,1]$. In *Riemann's* theory of integration, we partition the *domain* $[0,1]$ of the function into finitely many disjoint subintervals

$$[0,1] = \bigcup_{n=1}^{N} [x_{n-1}, x_n],$$

and denote the partition by $\mathcal{P} = \{0 = x_0 < x_1 < ... < x_N = 1\}$ and the length of the subinterval $[x_{n-1}, x_n]$ by $\triangle x_n = x_n - x_{n-1} > 0$. Then we define *upper and lower Riemann sums* associated with the partition $\mathcal{P}$ by

$$\begin{aligned} U(f;\mathcal{P}) &= \sum_{n=1}^{N} \left(\sup_{[x_{n-1},x_n]} f \right) \triangle x_n, \\ L(f;\mathcal{P}) &= \sum_{n=1}^{N} \left(\inf_{[x_{n-1},x_n]} f \right) \triangle x_n. \end{aligned}$$

Note that the suprema and infima are finite since f is bounded by assumption. Next we define the *upper and lower Riemann integrals* of f on $[0,1]$ by

$$\mathcal{U}(f) = \inf_{\mathcal{P}} U(f;\mathcal{P}), \qquad \mathcal{L}(f) = \sup_{\mathcal{P}} L(f;\mathcal{P}).$$

Thus the upper Riemann integral $\mathcal{U}(f)$ is the "smallest" of all the upper sums, and the lower Riemann integral is the "largest" of all the lower sums.

We can show that any upper sum is always larger than any lower sum by considering the *refinement* of two partitions $\mathcal{P}_1$ and $\mathcal{P}_2$: $\mathcal{P}_1 \cup \mathcal{P}_2$ denotes the paritition whose points consist of the union of the points in $\mathcal{P}_1$ and $\mathcal{P}_2$ and ordered to be strictly increasing.

Lemma 18. *Suppose $f : [0,1] \to \mathbb{R}$ is bounded. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are any two partitions of $[0,1]$, then*

$$U(f;\mathcal{P}_1) \geq U(f;\mathcal{P}_1 \cup \mathcal{P}_2) \geq L(f;\mathcal{P}_1 \cup \mathcal{P}_2) \geq L(f;\mathcal{P}_2). \tag{0.8}$$

Proof: Let

$$\begin{aligned} \mathcal{P}_1 &= \{0 = x_0 < x_1 < ... < x_M = 1\}, \\ \mathcal{P}_2 &= \{0 = y_0 < y_1 < ... < y_N = 1\}, \\ \mathcal{P}_1 \cup \mathcal{P}_2 &= \{0 = z_0 < z_1 < ... < z_P = 1\}. \end{aligned}$$

Fix a subinterval $[x_{n-1}, x_n]$ of the partition $\mathcal{P}_1$. Suppose that $[x_{n-1}, x_n]$ contains exactly the following increasing sequence of points in the partition $\mathcal{P}_1 \cup \mathcal{P}_2$:

$$z_{\ell_n} < z_{\ell_n+1} < ... < z_{\ell_n+m_n},$$

i.e. $z_{\ell_n} = x_{n-1}$ and $z_{\ell_n+m_n} = x_n$. Then we have

$$\begin{aligned}\left(\sup_{[x_{n-1},x_n]} f\right) \triangle x_n &= \left(\sup_{[x_{n-1},x_n]} f\right)\left(\sum_{j=1}^{m_n} \triangle z_{\ell_n+j}\right) \\ &\geq \sum_{j=1}^{m_n}\left(\sup_{[z_{\ell_n+j-1},z_{\ell_n+j}]} f\right) \triangle z_{\ell_n+j},\end{aligned}$$

since $\sup_{[z_{\ell_n+j-1},z_{\ell_n+j}]} f \leq \sup_{[x_{n-1},x_n]} f$ when $[z_{\ell_n+j-1}, z_{\ell_n+j}] \subset [x_{n-1}, x_n]$. If we now sum over $1 \leq n \leq M$ we get

$$\begin{aligned}U(f;\mathcal{P}_1) &= \sum_{n=1}^{M}\left(\sup_{[x_{n-1},x_n]} f\right) \triangle x_n \\ &\geq \sum_{n=1}^{M}\sum_{j=1}^{m_n}\left(\sup_{[z_{\ell_n+j-1},z_{\ell_n+j}]} f\right) \triangle z_{\ell_n+j} \\ &= \sum_{p=1}^{P}\left(\sup_{[z_{p-1},z_p]} f\right) \triangle z_p = U(f;\mathcal{P}_1 \cup \mathcal{P}_2).\end{aligned}$$

Similarly we can prove that

$$L(f;\mathcal{P}_2) \leq L(f;\mathcal{P}_1 \cup \mathcal{P}_2).$$

Since we trivially have $L(f;\mathcal{P}_1 \cup \mathcal{P}_2) \leq U(f;\mathcal{P}_1 \cup \mathcal{P}_2)$, the proof of the lemma is complete.

Now in (0.8) take the infimum over $\mathcal{P}_1$ and the supremum over $\mathcal{P}_2$ to obtain that

$$\mathcal{U}(f) \geq \mathcal{L}(f),$$

which says that the *upper* Riemann integral of f is always equal to or greater than the *lower* Riemann integral of f. Finally we say that f is Riemann integrable on $[0,1]$, written $f \in \mathcal{R}[0,1]$, if $\mathcal{U}(f) = \mathcal{L}(f)$, and we denote the common value by $\int_0^1 f$ or $\int_0^1 f(x)\,dx$.

We can of course repeat this line of definition and reasoning for any bounded closed interval $[a,b]$ in place of the closed unit interval $[0,1]$. We summarize matters in the following definition.

Definition 31. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. For any partition $\mathcal{P} = \{a = x_0 < x_1 < ... < x_N = b\}$ of $[a,b]$ we define* upper and lower Riemann sums *by*

$$\begin{aligned}U(f;\mathcal{P}) &= \sum_{n=1}^{N}\left(\sup_{[x_{n-1},x_n]} f\right) \triangle x_n, \\ L(f;\mathcal{P}) &= \sum_{n=1}^{N}\left(\inf_{[x_{n-1},x_n]} f\right) \triangle x_n.\end{aligned}$$

Set

$$\mathcal{U}(f) = \inf_{\mathcal{P}} U(f;\mathcal{P}), \qquad \mathcal{L}(f) = \sup_{\mathcal{P}} L(f;\mathcal{P}),$$

where the infimum and supremum are taken over all partitions $\mathcal{P}$ of $[a,b]$. We say that f is Riemann integrable *on $[a,b]$, written $f \in \mathcal{R}[a,b]$, if $\mathcal{U}(f) = \mathcal{L}(f)$, and we denote the common value by*

$$\int_a^b f \quad \text{or} \quad \int_a^b f(x)\,dx.$$

A more substantial generalization of the line of definition and reasoning above can be obtained on a closed interval $[a,b]$ by considering in place of the positive quantities $\triangle x_n = x_n - x_{n-1}$ associated with a partition

$$\mathcal{P} = \{a = x_0 < x_1 < ... < x_N = b\}$$

of $[a,b]$, the more general nonnegative quantities

$$\triangle\alpha_n = \alpha(x_n) - \alpha(x_{n-1}), \qquad 1 \le n \le N,$$

where $\alpha : [a,b] \to \mathbb{R}$ is *nondecreasing*. This leads to the notion of the Riemann-Stieltjes integral associated with a nondecreasing function $\alpha : [a,b] \to \mathbb{R}$.

Definition 32. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function and suppose $\alpha : [a,b] \to \mathbb{R}$ is nondecreasing. For any partition $\mathcal{P} = \{a = x_0 < x_1 < ... < x_N = b\}$ of $[a,b]$ we define* upper and lower Riemann sums *by*

$$\begin{aligned} U(f;\mathcal{P},\alpha) &= \sum_{n=1}^{N} \left(\sup_{[x_{n-1},x_n]} f \right) \triangle \alpha_n, \\ L(f;\mathcal{P},\alpha) &= \sum_{n=1}^{N} \left(\inf_{[x_{n-1},x_n]} f \right) \triangle \alpha_n. \end{aligned}$$

Set

$$\mathcal{U}(f,\alpha) = \inf_{\mathcal{P}} U(f;\mathcal{P},\alpha), \qquad \mathcal{L}(f,\alpha) = \sup_{\mathcal{P}} L(f;\mathcal{P},\alpha),$$

where the infimum and supremum are taken over all partitions $\mathcal{P}$ of $[a,b]$. We say that f is Riemann-Stieltjes integrable *on $[a,b]$, written $f \in \mathcal{R}_\alpha[a,b]$, if $\mathcal{U}(f,\alpha) = \mathcal{L}(f,\alpha)$, and we denote the common value by*

$$\int_a^b f d\alpha \quad \text{or} \quad \int_a^b f(x)\,d\alpha(x).$$

The lemma on partitions above generalizes immediately to the setting of the Riemann-Stieltjes integral.

Lemma 19. *Suppose $f : [a,b] \to \mathbb{R}$ is bounded and $\alpha : [a,b] \to \mathbb{R}$ is nondecreasing. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are any two partitions of $[a,b]$, then*

$$U(f;\mathcal{P}_1,\alpha) \ge U(f;\mathcal{P}_1 \cup \mathcal{P}_2,\alpha) \ge L(f;\mathcal{P}_1 \cup \mathcal{P}_2,\alpha) \ge L(f;\mathcal{P}_2,\alpha). \tag{0.9}$$

0.3. Existence of the Riemann-Stieltjes integral. The difficult question now arises as to exactly which bounded functions f are Riemann-Stieltjes integrable with respect to a given nondecreasing α on $[a,b]$. We will content ourselves with showing two results. Suppose f is bounded on $[a,b]$ and α is nondecreasing on $[a,b]$. Then

- $f \in \mathcal{R}_\alpha[a,b]$ if in addition f is *continuous* on $[a,b]$;
- $f \in \mathcal{R}_\alpha[a,b]$ if in addition f is *monotonic* on $[a,b]$ and α is *continuous* on $[a,b]$.

Both proofs will use the Cauchy criterion for existence of the integral $\int_a^b f d\alpha$ when $f : [a,b] \to \mathbb{R}$ is bounded and $\alpha : [a,b] \to \mathbb{R}$ is nondecreasing:

$$\text{For every } \varepsilon > 0 \text{ there is a partition } \mathcal{P} \text{ of } [a,b] \text{ such that} \quad U(f;\mathcal{P},\alpha) - L(f;\mathcal{P},\alpha) < \varepsilon. \tag{0.10}$$

Clearly, if (0.10) holds, then from (0.9) we obtain that for each $\varepsilon > 0$ that there is a partition $\mathcal{P}_\varepsilon$ satisfying

$$\begin{aligned} \mathcal{U}(f,\alpha) - \mathcal{L}(f,\alpha) &= \inf_{\mathcal{P}} U(f;\mathcal{P},\alpha) - \sup_{\mathcal{P}} L(f;\mathcal{P},\alpha) \\ &\leq U(f;\mathcal{P}_\varepsilon,\alpha) - L(f;\mathcal{P}_\varepsilon,\alpha) < \varepsilon. \end{aligned}$$

It follows that $\mathcal{U}(f,\alpha) = \mathcal{L}(f,\alpha)$ and so $\int_a^b f d\alpha$ exists. Conversely, given $\varepsilon > 0$ there are partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ satisfying

$$\begin{aligned} \mathcal{U}(f,\alpha) &= \inf_{\mathcal{P}} U(f;\mathcal{P},\alpha) > U(f;\mathcal{P}_1,\alpha) - \frac{\varepsilon}{2}, \\ \mathcal{L}(f,\alpha) &= \sup_{\mathcal{P}} L(f;\mathcal{P},\alpha) < L(f;\mathcal{P}_2,\alpha) + \frac{\varepsilon}{2}. \end{aligned}$$

Inequality (0.9) now shows that

$$\begin{aligned} U(f;\mathcal{P}_1 \cup \mathcal{P}_2,\alpha) - L(f;\mathcal{P}_1 \cup \mathcal{P}_2,\alpha) &\leq U(f;\mathcal{P}_1,\alpha) - L(f;\mathcal{P}_2,\alpha) \\ &< \left(\mathcal{U}(f,\alpha) + \frac{\varepsilon}{2}\right) - \left(\mathcal{L}(f,\alpha) - \frac{\varepsilon}{2}\right) = \varepsilon \end{aligned}$$

since $\mathcal{U}(f,\alpha) = \mathcal{L}(f,\alpha)$ if $\int_a^b f d\alpha$ exists. Thus we can take $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$ in (0.10).

The existence of $\int_a^b f d\alpha$ when f is continuous will use Theorem 30 on uniform continuity in a crucial way.

THEOREM 41. *Suppose that $f : [a,b] \to \mathbb{R}$ is continuous and $\alpha : [a,b] \to \mathbb{R}$ is nondecreasing. Then $f \in \mathcal{R}_\alpha[a,b]$.*

Proof: We will show that the Cauchy criterion (0.10) holds. Fix $\varepsilon > 0$. By Theorem 30 f is uniformly continuous on the compact set $[a,b]$, so there is $\delta > 0$ such that

$$|f(x) - f(x')| \leq \frac{\varepsilon}{\alpha(b) - \alpha(a)} \text{ whenever } |x - x'| \leq \delta.$$

Let $\mathcal{P} = \{a = x_0 < x_1 < ... < x_N = b\}$ be any partition of $[a,b]$ for which

$$\max_{1 \leq n \leq N} \triangle x_n < \delta.$$

Then we have

$$\sup_{[x_{n-1},x_n]} f - \inf_{[x_{n-1},x_n]} f \leq \sup_{x,x' \in [x_{n-1},x_n]} |f(x) - f(x')| \leq \varepsilon,$$

since $|x - x'| \leq \triangle x_n < \delta$ when $x, x' \in [x_{n-1}, x_n]$ by our choice of $\mathcal{P}$. Now we compute that

$$\begin{aligned} U(f;\mathcal{P},\alpha) - L(f;\mathcal{P},\alpha) &= \sum_{n=1}^N \left(\sup_{[x_{n-1},x_n]} f - \inf_{[x_{n-1},x_n]} f \right) \triangle \alpha_n \\ &\leq \sum_{n=1}^N \left(\frac{\varepsilon}{\alpha(b) - \alpha(a)} \right) \triangle \alpha_n = \varepsilon, \end{aligned}$$

which is (0.10) as required.

REMARK 17. *Observe that it makes no logical difference if we replace strict inequality* $<$ *with* $\leq$ *in '*$\varepsilon - \delta$ *type' definitions. We have used this observation twice in the above proof, and will continue to use it without further comment in the sequel.*

The proof of the next existence result uses the intermediate value theorem for continuous functions.

THEOREM 42. *Suppose that* $f : [a, b] \to \mathbb{R}$ *is monotone and* $\alpha : [a, b] \to \mathbb{R}$ *is nondecreasing and continuous. Then* $f \in \mathcal{R}_\alpha [a, b]$.

Proof: We will show that the Cauchy criterion (0.10) holds. Fix $\varepsilon > 0$ and suppose without loss of generality that f is nondecreasing on $[a, b]$. Let $N \geq 2$ be a positive integer. Since α is continuous we can use the intermediate value theorem to find points $x_n \in (a, b)$ such that $x_0 = a$, $x_N = b$ and

$$\alpha (x_n) = \alpha (a) + \frac{n}{N} (\alpha (b) - \alpha (a)), \qquad 1 \leq n \leq N - 1.$$

Since α is nondecreasing we have $x_{n-1} < x_n$ for all $1 \leq n \leq N$, and it follows that

$$\mathcal{P} = \{a = x_0 < x_1 < ... < x_N = b\}$$

is a partition of $[a, b]$ satisfying

$$\triangle\alpha_n = \alpha (x_n) - \alpha (x_{n-1}) = \frac{\alpha (b) - \alpha (a)}{N} < \frac{\varepsilon}{f (b) - f (a)},$$

provided we take N large enough. With such a partition $\mathcal{P}$ we compute

$$\begin{aligned} U (f; \mathcal{P}, \alpha) - L (f; \mathcal{P}, \alpha) &= \sum_{n=1}^{N} \left(\sup_{[x_{n-1}, x_n]} f - \inf_{[x_{n-1}, x_n]} f \right) \triangle \alpha_n \\ &\leq \frac{\varepsilon}{f (b) - f (a)} \sum_{n=1}^{N} \left(\sup_{[x_{n-1}, x_n]} f - \inf_{[x_{n-1}, x_n]} f \right) \\ &= \frac{\varepsilon}{f (b) - f (a)} \sum_{n=1}^{N} (f (x_n) - f (x_{n-1})) = \varepsilon, \end{aligned}$$

This proves (0.10) as required.

0.4. A stronger form of the definition of the Riemann integral. For the Riemann integral there is another formulation of the definition of $\int_a^b f$ that appears at first sight to be much stronger (and which doesn't work for general nondecreasing α in the Riemann-Stieltjes integral). For any partition $\mathcal{P} = \{a = x_0 < x_1 < ... < x_N = b\}$, set $\|\mathcal{P}\| = \max_{1 \leq n \leq N} \triangle x_n$, called the *norm* of $\mathcal{P}$. Now if $\int_a^b f$ exists, then for every $\varepsilon > 0$ there is by the Cauchy criterion (0.10) a partition $\mathcal{P} = \{a = x_0 < x_1 < ... < x_N = b\}$ such that

$$U (f; \mathcal{P}) - L (f; \mathcal{P}) < \frac{\varepsilon}{2}.$$

Now define δ to be the smaller of the two positive numbers

$$\min_{1 \leq n \leq N} \triangle x_n \text{ and } \frac{\varepsilon}{2N \ diam \ f ([a, b])}.$$

CLAIM 1. *If* $\mathcal{Q} = \{a = y_0 < y_1 < ... < y_M = b\}$ *is* any *partition with*

$$\|\mathcal{Q}\| = \max_{1 \leq m \leq M} \triangle y_m < \delta,$$

then

$$U(f;\mathcal{Q}) - L(f;\mathcal{Q}) < \varepsilon.$$

Indeed, since $\triangle y_m < \delta \le \triangle x_n$ for all m and n by choice of δ, each point x_n lies in a *distinct* one of the subintervals $[y_{m-1}, y_m]$ of $\mathcal{Q}$, call it $J_n = [y_{m_n-1}, y_{m_n}]$. The other subintervals $[y_{m-1}, y_m]$ of $\mathcal{Q}$ with m not equal to any of the m_n, each lie in one of the separating intervals $K_n = \left[y_{m_{n-1}}, y_{m_n-1}\right]$ that are formed by the spaces between the intervals J_n. These intervals K_n are the union of one or more consecutive subintervals of $\mathcal{Q}$. We have for each n that

$$\begin{aligned}
&\sum_{m:[y_{m-1},y_m]\subset K_n} \left(\sup_{[y_{m-1},y_m]} f - \inf_{[y_{m-1},y_m]} f \right) \triangle y_m \\
&\le \left(\sup_{\left[y_{m_{n-1}},y_{m_n-1}\right]} f - \inf_{\left[y_{m_{n-1}},y_{m_n-1}\right]} f \right) \sum_{m:[y_{m-1},y_m]\subset K_n} \triangle y_m \\
&\le \left(\sup_{[x_n,x_{n+1}]} f - \inf_{[x_n,x_{n+1}]} f \right) (y_m - y_{m-1}) \\
&\le \left(\sup_{[x_n,x_{n+1}]} f - \inf_{[x_n,x_{n+1}]} f \right) (x_{n+1} - x_n).
\end{aligned}$$

Summing this in n yields

$$\begin{aligned}
&\sum_{n=1}^{N} \sum_{m:[y_{m-1},y_m]\subset K_n} \left(\sup_{[y_{m-1},y_m]} f - \inf_{[y_{m-1},y_m]} f \right) \triangle y_m \\
\le\ &\sum_{n=1}^{N} \left(\sup_{[x_n,x_{n+1}]} f - \inf_{[x_n,x_{n+1}]} f \right) (x_{n+1} - x_n) = U(f;\mathcal{P}) - L(f;\mathcal{P}).
\end{aligned} \tag{0.11}$$

Now we compute

$$\begin{aligned}
U(f;\mathcal{Q}) - L(f;\mathcal{Q}) &= \sum_{m=1}^{M} \left(\sup_{[y_{m-1},y_m]} f - \inf_{[y_{m-1},y_m]} f \right) \triangle y_m \\
&= \sum_{n=1}^{N} \left(\sup_{J_n} f - \inf_{J_n} f \right) (y_{m_n} - y_{m_n-1}) \\
&\quad + \sum_{n=1}^{N} \sum_{m:[y_{m-1},y_m]\subset K_n} \left(\sup_{[y_{m-1},y_m]} f - \inf_{[y_{m-1},y_m]} f \right) \triangle y_m,
\end{aligned}$$

which by (0.11) and choice of δ is dominated by

$$\begin{aligned}
&\mathit{diam}\ f([a,b]) \sum_{n=1}^{N} (y_{m_n} - y_{m_n-1}) + U(f;\mathcal{P}) - L(f;\mathcal{P}) \\
\le\ &\mathit{diam}\ f([a,b])\, N\delta + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,
\end{aligned}$$

and this proves the claim.

Conversely, if

$$\begin{aligned}
&\text{For every } \varepsilon > 0 \text{ there is } \delta > 0 \text{ such that} \\
&\qquad U(f;\mathcal{Q}) - L(f;\mathcal{Q}) < \varepsilon \text{ whenever } \|\mathcal{Q}\| < \delta,
\end{aligned} \tag{0.12}$$

then the Cauchy criterion (0.10) holds with $\mathcal{P}$ equal to any such $\mathcal{Q}$. Thus (0.12) provides another equivalent definition of the Riemann integral $\int_a^b f$ that is more like the $\varepsilon - \delta$ definition of continuity at a point (compare Definition 25).

1. Simple properties of the Riemann-Stieltjes integral

The Riemann-Stieltjes integral $\int_a^b f d\alpha$ is a function of the closed interval $[a, b]$, the bounded function f on $[a, b]$, and the nondecreasing function α on $[a, b]$. With respect to each of these three variables, the integral has natural properties related to monotonicity, sums and scalar multiplication. In fact we have the following lemmas dealing with each variable separately, beginning with f, then α and ending with $[a, b]$.

LEMMA 20. *Fix $[a, b] \subset \mathbb{R}$ and $\alpha : [a, b] \to \mathbb{R}$ nondecreasing. The set $\mathcal{R}_\alpha [a, b]$ is a real vector space and the integral $\int_a^b f d\alpha$ is a* linear *function of $f \in \mathcal{R}_\alpha [a, b]$: if $f_j \in \mathcal{R} [a, b]$ and $\lambda_j \in \mathbb{R}$, then*

$$f = \lambda_1 f_1 + \lambda_2 f_2 \in \mathcal{R}_\alpha [a, b] \ \ \text{and} \ \int_a^b f d\alpha = \lambda_1 \int_a^b f_1 d\alpha + \lambda_2 \int_a^b f_2 d\alpha.$$

Furthermore, $\mathcal{R}_\alpha [a, b]$ is partially ordered by declaring $f \leq g$ if $f(x) \leq g(x)$ for $x \in [a, b]$, and the integral $\int_a^b f d\alpha$ is a nondecreasing *function of f with respect to this order: if $f, g \in \mathcal{R}_\alpha [a, b]$ and $f \leq g$, then $\int_a^b f d\alpha \leq \int_a^b g d\alpha$.*

LEMMA 21. *Fix $[a, b] \subset \mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ bounded. Then*

$$\mathcal{C}_f [a, b] \equiv \{\alpha : [a, b] \to \mathbb{R} : \alpha \text{ is nondecreasing and } f \in \mathcal{R}_\alpha [a, b]\}$$

is a cone and the integral $\int_a^b f d\alpha$ is a 'positive linear' *function of α: if $\alpha_j \in \mathcal{C}_f [a, b]$ and $c_j \in [0, \infty)$, then*

$$\alpha = c_1 \alpha_1 + c_2 \alpha_2 \in \mathcal{C}_f [a, b] \ \ \text{and} \ \int_a^b f d\alpha = c_1 \int_a^b f d\alpha_1 + c_2 \int_a^b f d\alpha_2.$$

LEMMA 22. *Fix $[a, b] \subset \mathbb{R}$ and $\alpha : [a, b] \to \mathbb{R}$ nondecreasing and $f \in \mathcal{R}_\alpha [a, b]$. If $a < c < b$, then $\alpha : [a, c] \to \mathbb{R}$ and $\alpha : [c, b] \to \mathbb{R}$ are each nondecreasing and*

$$f \in \mathcal{R}_\alpha [a, c] \ \ \text{and} \ f \in \mathcal{R}_\alpha [c, b] \ \ \text{and} \ \int_a^b f d\alpha = \int_a^c f d\alpha + \int_c^b f d\alpha.$$

These three lemmas are easy to prove, and are left to the reader. Properties regarding multiplication of functions in $\mathcal{R}_\alpha [a, c]$ and composition of functions are more delicate.

THEOREM 43. *Suppose that $f : [a, b] \to [m, M]$ and $f \in \mathcal{R}_\alpha [a, b]$. If $\varphi : [m, M] \to \mathbb{R}$ is continuous, then $\varphi \circ f \in \mathcal{R}_\alpha [a, b]$.*

COROLLARY 13. *If $f, g \in \mathcal{R}_\alpha [a, b]$, then $fg \in \mathcal{R}_\alpha [a, b]$, $|f| \in \mathcal{R}_\alpha [a, b]$ and*

$$\left| \int_a^b f d\alpha \right| \leq \int_a^b |f| \, d\alpha.$$

Proof: Since $\varphi(x) = x^2$ is continuous, Lemma 20 and Theorem 43 yield

$$fg = \frac{1}{2} \left\{ (f + g)^2 - f^2 - g^2 \right\} \in \mathcal{R}_\alpha [a, b] \, .$$

Since $\varphi(x) = |x|$ is continuous, Theorem 43 yields $|f| \in \mathcal{R}_\alpha[a,b]$. Now choose $c = \pm 1$ so that $c\int_a^b f d\alpha \geq 0$. Then the lemmas imply

$$\left|\int_a^b f d\alpha\right| = c\int_a^b f d\alpha = \int_a^b (cf)\, d\alpha \leq \int_a^b |f|\, d\alpha.$$

Proof (of Theorem 43): Let $h = \varphi \circ f$. We will show that $h \in \mathcal{R}_\alpha[a,b]$ by verifying the Cauchy criterion for integrals (0.10). Fix $\varepsilon > 0$. Since φ is continuous on the compact interval $[m,M]$, it is uniformly continuous on $[m,M]$ by Theorem 30. Thus we can choose $0 < \delta < \varepsilon$ such that

$$|\varphi(s) - \varphi(t)| < \varepsilon \text{ whenever } |s-t| < \delta.$$

Since $f \in \mathcal{R}_\alpha[a,b]$, there is by the Cauchy criterion a partition

$$\mathcal{P} = \{a = x_0 < x_1 < ... < x_N = b\}$$

such that

$$U(f;\mathcal{P},\alpha) - L(f;\mathcal{P},\alpha) < \delta^2. \tag{1.1}$$

Let

$$\begin{aligned} M_n &= \sup_{[x_{n-1},x_n]} f \text{ and } m_n = \inf_{[x_{n-1},x_n]} f, \\ M_n^* &= \sup_{[x_{n-1},x_n]} h \text{ and } m_n^* = \inf_{[x_{n-1},x_n]} h, \end{aligned}$$

and set

$$A = \{n : M_n - m_n < \delta\} \text{ and } A = \{n : M_n - m_n \geq \delta\}.$$

The point of the index set A is that for each $n \in A$ we have

$$\begin{aligned} M_n^* - m_n^* &= \sup_{x,y\in[x_{n-1},x_n]} |\varphi(f(x)) - \varphi(f(y))| \leq \sup_{|s-t|\leq M_n - m_n} |\varphi(s) - \varphi(t)| \\ &\leq \sup_{|s-t|<\delta} |\varphi(s) - \varphi(t)| \leq \varepsilon, \qquad n \in A. \end{aligned}$$

As for n in the index set B, we have $\triangle\alpha_n \leq M_n - m_n$ and the inequality (1.1) then gives

$$\delta \sum_{n\in B} \triangle\alpha_n \leq \sum_{n\in B} (M_n - m_n) \triangle \alpha_n < \delta^2.$$

Dividing by $\delta > 0$ we obtain

$$\sum_{n\in B} \triangle\alpha_n < \delta.$$

Now we use the trivial bound

$$M_n^* - m_n^* \leq diam\ \varphi([m,M])$$

to compute that

$$\begin{aligned} U(h;\mathcal{P},\alpha) - L(h;\mathcal{P},\alpha) &= \left\{\sum_{n\in A} + \sum_{n\in B}\right\} (M_n^* - m_n^*) \triangle \alpha_n \\ &\leq \sum_{n\in A} \varepsilon \triangle \alpha_n + \sum_{n\in B} diam\ \varphi([m,M]) \triangle \alpha_n \\ &\leq \varepsilon(\alpha(b) - \alpha(a)) + \delta\ diam\ \varphi([m,M]) \\ &\leq \varepsilon[\alpha(b) - \alpha(a) + diam\ \varphi([m,M])], \end{aligned}$$

which verifies (0.10) for the existence of $\int_a^b h d\alpha$ as required.

1.1. The Henstock-Kurtzweil integral. We can reformulate the $\varepsilon-\delta$ definition of the Riemann integral $\int_a^b f$ in (0.12) using a more general notion of partition, that of a tagged partition. If $\mathcal{P} = \{a = x_0 < x_1 < ... < x_N = b\}$ is a partition of $[a,b]$ and we choose points $t_n \in [x_{n-1}, x_n]$ in each subinterval of $\mathcal{P}$, then

$$\mathcal{P}^* = \{a = x_0 \leq t_1 \leq x_1 \leq ... \leq x_{N-1} \leq t_N \leq x_N = b\}, \quad \text{where } x_0 < x_1 < ... < x_N,$$

is called a *tagged* partition $\mathcal{P}^*$ with *underlying* partition $\mathcal{P}$. Thus a tagged partition consists of two finite intertwined sequences $\{x_n\}_{n=0}^N$ and $\{t_n\}_{n=1}^N$, where the sequence $\{x_n\}_{n=0}^N$ is strictly increasing and the sequence $\{t_n\}_{n=1}^N$ need not be. For every tagged partition $\mathcal{P}^*$ of $[a,b]$, define the corresponding *Riemann sum* $S(f;\mathcal{P}^*)$ by

$$S(f;\mathcal{P}^*) = \sum_{n=1}^N f(t_n) \triangle x_n.$$

Note that $\inf_{[x_{n-1},x_n]} f \leq f(t_n) \leq \sup_{[x_{n-1},x_n]} f$ implies that

$$L(f;\mathcal{P}) \leq S(f;\mathcal{P}^*) \leq U(f;\mathcal{P})$$

for all tagged partitions $\mathcal{P}^*$ with underlying partition $\mathcal{P}$.

Now observe that if $f \in \mathcal{R}[a,b]$, $\varepsilon > 0$ and the partition $\mathcal{P}$ satisfies

$$U(f;\mathcal{P}) - L(f;\mathcal{P}) < \varepsilon,$$

then *every* tagged partition $\mathcal{P}^*$ with underlying partition $\mathcal{P}$ satisfies

$$\left| S(f;\mathcal{P}^*) - \int_a^b f \right| \leq U(f;\mathcal{P}) - L(f;\mathcal{P}) < \varepsilon. \tag{1.2}$$

Conversely if for each $\varepsilon > 0$ there is a partition $\mathcal{P}$ such that *every* tagged partition $\mathcal{P}^*$ with underlying partition $\mathcal{P}$ satisfies (1.2), then (0.10) holds and so $f \in \mathcal{R}[a,b]$.

However, we can also formulate this approach using the $\varepsilon - \delta$ form (0.12) of the definition of $\int_a^b f$. The result is that $f \in \mathcal{R}[a,b]$ if and only if

(1.3) There is $L \in \mathbb{R}$ such that for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$|S(f;\mathcal{P}^*) - L| < \varepsilon \text{ whenever } \|\mathcal{P}^*\| < \delta.$$

Of course if such a number L exists we write $L = \int_a^b f$ and call it the Riemann integral of f on $[a,b]$. Here we define $\|\mathcal{P}^*\|$ to be $\|\mathcal{P}\|$ where $\mathcal{P}$ is the underlying partition of $\mathcal{P}^*$. The reader can easily verify that $f \in \mathcal{R}[a,b]$ if and only if the above condition (1.3) holds.

Now comes the clever insight of Henstock and Kurtzweil. We view the positive constant δ in (1.3) as a *function* on the interval $[a,b]$, and replace it with an arbitrary (not necessarily constant) positive function $\delta : [a,b] \to (0,\infty)$. We refer to such an arbitrary positive function $\delta : [a,b] \to (0,\infty)$ as a *guage* on $[a,b]$. Then for any guage on $[a,b]$, we say that a tagged partition $\mathcal{P}^*$ on $[a,b]$ is *δ-fine* provided

$$[x_{n-1}, x_n] \subset (t_n - \delta(t_n), t_n + \delta(t_n)), \qquad 1 \leq n \leq N. \tag{1.4}$$

Thus $\mathcal{P}^*$ is δ-fine if each tag $t_n \in [x_{n-1}, x_n]$ has its associated guage value $\delta(t_n)$ sufficiently large that the open interval centered at t_n with radius $\delta(t_n)$ *contains* the n^{th} subinterval $[x_{n-1}, x_n]$ of the partition $\mathcal{P}$. Now we can give the definition of the Henstock and Kurtzweil integral.

DEFINITION 33. *A function $f : [a, b] \to \mathbb{R}$ is Henstock-Kurtzweil integrable on $[a, b]$, written $f \in \mathcal{HK}[a, b]$, if there is $L \in \mathbb{R}$ such that for every $\varepsilon > 0$ there is a guage $\delta_\varepsilon : [a, b] \to (0, \infty)$ on $[a, b]$ such that*

$$|S(f; \mathcal{P}^*) - L| < \varepsilon \text{ whenever } \mathcal{P}^* \text{ is } \delta\text{-fine.}$$

It is clear that if $f \in \mathcal{R}[a, b]$ is Riemann integrable, then f satisfies Definition 33 with $L = \int_a^b f$ - simply take δ_ε to be the constant guage δ in (1.3). However, for this new definition to have any value it is necessary that such an L is uniquely determined by Definition 33. This is indeed the case and relies crucially on the fact that $[a, b]$ is compact. Here are the details.

Suppose that Definition 33 holds with both L and L'. Let $\varepsilon > 0$. Then there are guages δ_ε and δ'_ε on $[a, b]$ such that

$$\begin{aligned} |S(f; \mathcal{P}^*) - L| &< \varepsilon \text{ whenever } \mathcal{P}^* \text{ is } \delta_\varepsilon\text{-fine,} \\ |S(f; \mathcal{P}^*) - L'| &< \varepsilon \text{ whenever } \mathcal{P}^* \text{ is } \delta'_\varepsilon\text{-fine.} \end{aligned}$$

Now define

$$\eta_\varepsilon(x) = \min\left\{\delta_\varepsilon(x), \delta'_\varepsilon(x)\right\}, \qquad a \le x \le b.$$

Then η_ε is a guage on $[a, b]$. Here is the critical point: we would like to produce a tagged partition $\mathcal{P}^*_\varepsilon$ that is η_ε-fine! Indeed, if such a tagged partition $\mathcal{P}^*_\varepsilon$ exists, then $\mathcal{P}^*_\varepsilon$ would also be δ_ε-fine and δ'_ε-fine (since $\eta_\varepsilon \le \delta_\varepsilon$ and $\eta_\varepsilon \le \delta'_\varepsilon$) and hence

$$|L - L'| \le |S(f; \mathcal{P}^*_\varepsilon) - L| + |S(f; \mathcal{P}^*_\varepsilon) - L'| < 2\varepsilon$$

for all $\varepsilon > 0$, which forces $L = L'$.

However, if η is any guage on $[a, b]$, let

$$B(x, \eta(x)) = (x - \eta(x), x + \eta(x)) \text{ and } B\left(x, \frac{\eta(x)}{2}\right) = \left(x - \frac{\eta(x)}{2}, x + \frac{\eta(x)}{2}\right).$$

Then $\left\{B\left(x, \frac{\eta(x)}{2}\right)\right\}_{x \in [a,b]}$ is an open cover of the compact set $[a, b]$, hence there is a finite subcover $\left\{B\left(x_n, \frac{\eta(x_n)}{2}\right)\right\}_{n=0}^N$. We may assume that every interval $B\left(x_n, \frac{\eta(x_n)}{2}\right)$ is needed to cover $[a, b]$ by discarding any in turn which are included in the union of the others. We may also assume that $a = x_0 < x_1 < ... < x_N = b$. It follows that $B\left(x_{n-1}, \frac{\eta(x_{n-1})}{2}\right) \cap B\left(x_n, \frac{\eta(x_n)}{2}\right) \neq \emptyset$, so the triangle inequality yields

$$|x_n - x_{n-1}| < \frac{\eta(x_{n-1}) + \eta(x_n)}{2}, \qquad 1 \le n \le N.$$

If $\eta(x_n) \ge \eta(x_{n-1})$ then

$$[x_{n-1}, x_n] \subset B(x_n, \eta(x_n)),$$

and so we define

$$t_n = x_n.$$

Otherwise, we have $\eta(x_{n-1}) > \eta(x_n)$ and then

$$[x_{n-1}, x_n] \subset B(x_{n-1}, \eta(x_{n-1})),$$

and so we define

$$t_n = x_{n-1}.$$

The tagged partition

$$\mathcal{P}^* = \{a = x_0 \leq t_1 \leq x_1 \leq ... \leq x_{N-1} \leq t_N \leq x_N = b\}$$

is then η-fine.

With the uniqueness of the Henstock-Kurtzweil integral in hand, and the fact that it extends the definition of the Riemann integral, we can without fear of confusion denote the Henstock-Kurtzweil integral by $\int_a^b f$ when $f \in \mathcal{HK}[a,b]$. It is now possible to develop the standard properties of these integrals as in Theorem 43 and the lemmas above for Riemann integrals. The proofs are typically very similar to those commonly used for Riemann integration. One exception is the Fundamental Theorem of Calculus for the Henstock-Kurtzweil integral, which requires a more complicated proof. In fact, it turns out that the theory of the Henstock-Kurtzweil integral is sufficiently rich to *include* the theory of the Lebesgue integral, which we consider in detail in a later chapter. For further development of the theory of the Henstock-Kurtzweil integral we refer the reader to Bartle and Sherbert [**1**] and the references given there.

2. Fundamental Theorem of Calculus

The operations of integration and differentiation are inverse to each other in a certain sense which we make precise in this section. We consider only the Riemann integral. Our first theorem proves a sense in which

$$Differentiation \circ Integration = Identity,$$

and the second theorem proves a sense in which

$$Integration \circ Differentiation = Identity.$$

The second theorem is often called the Fundamental Theorem of Calculus, while the two together are sometimes referred to in this way. As an application we derive an integration by parts formula in the third theorem below.

THEOREM 44. *Suppose $f \in \mathcal{R}[a,b]$. Define*

$$F(x) = \int_a^x f(t)\,dt, \qquad \text{for } a \leq x \leq b.$$

Then F is continuous on $[a,b]$ and

$$F'(x) \text{ exists and equals } f(x)$$

at every point $x \in [a,b]$ at which f is continuous.

Proof: First we show that F is continuous on $[a,b]$. Since f is bounded there is a positive M such that $|f(x)| \leq M$ for $a \leq x \leq b$. Then Lemma 22 yields

$$|F(y) - F(x)| = \left|\int_a^y f(t)\,dt - \int_a^x f(t)\,dt\right| = \left|\int_x^y f(t)\,dt\right|,$$

and if we apply Corollary 13 we obtain for $a \leq x < y \leq b$,

$$|F(x) - F(y)| \leq \int_x^y |f(t)|\,dt \leq \int_x^y M dt = M(y-x) = M|x-y|.$$

This easily gives the continuity of F on $[a,b]$, in fact it implies the uniform continuity of F on $[a,b]$: $|F(x)-F(y)|<\varepsilon$ whenever $|x-y|<\delta\equiv\frac{\varepsilon}{M}$.

Now suppose that f is continuous at a fixed $x_0\in[a,b]$. Given $\varepsilon>0$ choose $\delta>0$ so that

$$|f(x)-f(x_0)|<\varepsilon \text{ if } |x-x_0|<\delta \text{ and } x\in[a,b].$$

Then if $t\in(x_0,x_0+\delta)\cap[a,b]$ we have

$$\begin{aligned}\left|\frac{F(t)-F(x_0)}{t-x_0}-f(x_0)\right| &= \left|\frac{\int_{x_0}^t f(x)\,dx}{t-x_0}-f(x_0)\right| \\ &= \left|\frac{1}{t-x_0}\int_{x_0}^t [f(x)-f(x_0)]\,dx\right| \\ &\leq \frac{1}{t-x_0}\int_{x_0}^t |f(x)-f(x_0)|\,dx<\varepsilon.\end{aligned}$$

Similarly if $t\in(x_0-\delta,x_0)\cap[a,b]$ we have

$$\left|\frac{F(x_0)-F(t)}{x_0-t}-f(x_0)\right|<\varepsilon.$$

This proves that $\lim_{t\to x_0}\frac{F(x_0)-F(t)}{x_0-t}=f(x_0)$ as required.

THEOREM 45. *Suppose* $f\in\mathcal{R}[a,b]$. *If there is a continuous function* F *on* $[a,b]$ *that is differentiable on* (a,b) *and satisfies*

$$F'(x)=f(x),\qquad x\in(a,b),$$

then

$$\int_a^b f(x)\,dx=F(b)-F(a). \tag{2.1}$$

Proof: Given $\varepsilon>0$ use the Cauchy criterion for integrals (0.10) to choose a partition

$$\mathcal{P}=\{a=x_0<x_1<...<x_N=b\}$$

of $[a,b]$ satisfying

$$U(f;\mathcal{P})-L(f;\mathcal{P})<\varepsilon.$$

Now apply the second mean value Theorem 35 to F on the subinterval $[x_{n-1},x_n]$ to obtain points $t_n\in(x_{n-1},x_n)$ such that

$$F'(t_n)=\frac{F(x_n)-F(x_{n-1})}{x_n-x_{n-1}},$$

so that

$$F(x_n)-F(x_{n-1})=F'(t_n)\triangle x_n=f(t_n)\triangle x_n.$$

Thus we have

$$F(b)-F(a)=\sum_{n=0}^N(F(x_n)-F(x_{n-1}))=\sum_{n=0}^N f(t_n)\triangle x_n.$$

But (1.2) implies that

$$\left|\int_a^b f-\sum_{n=0}^N f(t_n)\triangle x_n\right|\leq U(f;\mathcal{P})-L(f;\mathcal{P})<\varepsilon,$$

and we conclude that

$$\left| F(b) - F(a) - \int_a^b f \right| < \varepsilon$$

for every $\varepsilon > 0$, hence (2.1) holds.

Theorem 46. *(Integration by parts) Suppose that F, G are differentiable functions on $[a,b]$ with $F', G' \in \mathcal{R}[a,b]$. Then*

$$\int_a^b F'(x) G(x)\, dx + \int_a^b F(x) G'(x)\, dx = F(b) G(b) - F(a) G(a).$$

Proof: By Proposition 15 the function $H(x) = F(x) G(x)$ has derivative

$$H'(x) = F'(x) G(x) + F(x) G'(x),$$

and by Lemma 17 and Theorems 41 and 13 we have

$$H' \in \mathcal{R}[a,b].$$

Now we apply (2.1) to H and $h = H'$ to obtain

$$H(b) - H(a) = \int_a^b h = \int_a^b (F'G + FG') = \int_a^b F'G + \int_a^b FG'.$$

CHAPTER 7

Function spaces

A very powerful abstract idea in analysis is to consider metric spaces whose *points* consist of *functions* defined on yet another metric space. A prime example is the 'metric space of functions' $C_{\mathbb{R}}(X)$, which we now define. Suppose X is a *compact* metric space and let

$$C_{\mathbb{R}}(X) = \{f : X \to \mathbb{R} : f \text{ is continuous}\},$$

be the set of all continuous functions f mapping X into the real numbers $\mathbb{R}$. Clearly $C_{\mathbb{R}}(X)$ is a real vector space with the usual notion of addition of functions and scalar multiplication. However, we can also define a metric structure on $C_{\mathbb{R}}(X)$ as follows. For $f, g \in C_{\mathbb{R}}(X)$, define

$$d(f,g) = d_{C_{\mathbb{R}}(X)}(f,g) = \sup_{x\in X} |f(x) - g(x)|. \tag{0.2}$$

Since $f - g \in C_{\mathbb{R}}(X)$ is continuous on a compact set X, and the absolute value function is continuous, it follows from Theorem 28 that the supremum defining $d(f,g)$ is a finite nonnegative real number (and is even achieved as $|f(x) - g(x)|$ for some $x \in X$). Note that in the case $X = [a,b]$ is a closed interval on the real line, the quantity $d(f,g)$ is the largest vertical distance between points on the graphs of f and g. It is an easy exercise to verify that $d : X \times X \to [0,\infty)$ satisfies the axioms of a metric. In particular, if $f, g, h \in C_{\mathbb{R}}(X)$, then

$$\begin{aligned} d(f,h) &= \sup_{x\in X} |f(x) - h(x)| = \sup_{x\in X} |[f(x) - g(x)] + [g(x) - h(x)]| \\ &\leq \sup_{x\in X} |f(x) - g(x)| + \sup_{x\in X} |g(x) - h(x)| = d(f,g) + d(g,h). \end{aligned}$$

Thus $(C_{\mathbb{R}}(X), d)$ is a metric space whose elements are continuous real-valued functions on X. The single most important result of this chapter is that this particular metric space is *complete*, i.e. every Cauchy sequence in $C_{\mathbb{R}}(X)$ converges. A crucial role is played here by an investigation of limits of sequences in $C_{\mathbb{R}}(X)$, namely limits of sequences of continuous functions on X.

1. Sequences and series of functions

We begin by examining more carefully the notion of convergence of a sequence of functions in the metric space $C_{\mathbb{R}}(X)$. We begin with a general definition of uniform convergence.

Definition 34. *Suppose X and Y are metric spaces and $E \subset X$. Suppose that $\{f_n\}_{n=1}^{\infty}$ is a sequence of functions $f_n : E \to Y$ and that $f : E \to Y$. We say that the sequence $\{f_n\}_{n=1}^{\infty}$ converges* uniformly *to f on E if for every $\varepsilon > 0$ there is a positive integer N such that*

$$d_Y(f_n(x), f(x)) \leq \varepsilon \text{ for all } n \geq N \text{ and all } x \in E. \tag{1.1}$$

In this case we write $f_n \to f$ uniformly on E.

Note in particular that if $f_n \to f$ uniformly on E then the sequence $\{f_n\}_{n=1}^{\infty}$ converges *pointwise* to f on E, written $f_n \to f$ pointwise on E, by which we mean

$$\lim_{n\to\infty} f_n(x) = f(x),$$

for each $x \in E$. The point of uniform convergence of the sequence $\{f_n\}_{n=1}^{\infty}$ is that there is a positive integer N that depends only on ε and *not* on $x \in E$, that works in (1.1).

EXAMPLE 9. *Let $f_n : [0,1] \to [0,1]$ by $f_n(x) = x^n$. Let*

$$f(x) = \begin{cases} 0 & \text{if} \quad 0 \le x < 1 \\ 1 & \text{if} \quad x = 1 \end{cases}.$$

Then $f_n \to f$ pointwise on $[0,1]$ but the convergence is not *uniform. Indeed, for any $n \ge 1$ there is a point $x \in [0,1)$ such that*

$$|f_n(x) - f(x)| = |x^n - 0| = x^n \ge \frac{1}{2}.$$

This is because the monomial x^n is continuous and so $\lim_{x\to 1} x^n = 1^n = 1$.

An important feature of this example is that the functions f_n are each continuous on the set $[0,1]$ (which also happens to be compact), yet their pointwise limit is *not* continuous on $[0,1]$. The next theorem shows that the reason can be attributed to the failure of uniform convergence here.

THEOREM 47. *Suppose that X and Y are metric spaces and $E \subset X$. Suppose also that $\{f_n\}_{n=1}^{\infty}$ is a sequence of* continuous *functions from E to Y and that $f : E \to Y$. If $f_n \to f$ uniformly on E, then f is continuous on E.*

Proof: Fix a point $p \in E$ and let $\varepsilon > 0$. We must show that there is $\delta > 0$ such that

$$d_Y(f(x), f(p)) < \varepsilon \text{ whenever } d_X(x,p) < \delta \text{ and } x \in E.$$

Since $f_n \to f$ uniformly on E we can choose N so large that (1.1) holds with $\frac{\varepsilon}{3}$ in place of ε:

$$d_Y(f_n(x), f(x)) \le \frac{\varepsilon}{3} \text{ for } \textit{all } n \ge N \text{ and } \textit{all } x \in E. \tag{1.2}$$

Now use the continuity of f_N on E at the point p to find $\delta > 0$ satisfying

$$d_Y(f_N(x), f_N(p)) < \frac{\varepsilon}{3} \text{ whenever } d_X(x,p) < \delta \text{ and } x \in E. \tag{1.3}$$

Finally the triangle inequality yields

$$\begin{aligned} d_Y(f(x), f(p)) &\le d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(p)) + d_Y(f_N(p), f(p)) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

whenever $d_X(x,p) < \delta$ and $x \in E$, upon applying (1.2) with $n = N$ to the first and third terms on the right, and applying (1.3) to the middle term on the right.

2. The metric space $C_{\mathbb{R}}(X)$

We can now prove the main result of this chapter, namely that the metric space $C_{\mathbb{R}}(X)$ is complete. Recall that X is compact now. The connection with uniform convergence is this: a sequence $\{f_n\}_{n=1}^{\infty}$ in $C_{\mathbb{R}}(X)$ converges to $f \in C_{\mathbb{R}}(X)$ in the metric d of $C_{\mathbb{R}}(X)$ given in (0.2), *if and only if* $f_n \to f$ uniformly on X. This is in fact a definition chaser as in the case $E = X$ and $Y = \mathbb{R}$, (1.1) says precisely that

$$d(f_n, f) = \sup_{x \in X} |f_n(x) - f(x)| \leq \varepsilon \text{ for all } n \geq N.$$

It follows immediately that $f_n \to f$ in $C_{\mathbb{R}}(X)$ if and only if $f_n \to f$ uniformly on X.

THEOREM 48. *Let X be a compact metric space. Then the metric space $C_{\mathbb{R}}(X)$ is complete.*

Proof: Let $\{f_n\}_{n=1}^{\infty}$ be a Cauchy sequence in $C_{\mathbb{R}}(X)$. We must show that $\{f_n\}_{n=1}^{\infty}$ converges to some $f \in C_{\mathbb{R}}(X)$. Now for every $\varepsilon > 0$ there is N such that

$$\sup_{x \in X} |f_m(x) - f_n(x)| = d(f_n, f_m) \leq \varepsilon \qquad \text{for all } m, n \geq N.$$

In particular for each $x \in X$ the sequence $\{f_n(x)\}_{n=1}^{\infty}$ is Cauchy in $\mathbb{R}$. Since the real numbers $\mathbb{R}$ are complete, there is for each $x \in X$ a real number $f(x)$ such that

$$\lim_{n \to \infty} f_n(x) = f(x).$$

Moreover for $m \geq N$ and $x \in X$ we have

$$|f_m(x) - f(x)| = \left| f_m(x) - \lim_{n \to \infty} f_n(x) \right| = \lim_{n \to \infty} |f_m(x) - f_n(x)| \leq \lim_{n \to \infty} \varepsilon = \varepsilon.$$

This shows that $f_m \to f$ uniformly on X. Now we apply Theorem 47 to conclude that f is continuous on X, i.e. $f \in C_{\mathbb{R}}(X)$. We've already noted that in the metric space $C_{\mathbb{R}}(X)$, $f_m \to f$ in $C_{\mathbb{R}}(X)$ is equivalent to $f_m \to f$ uniformly on X. Thus we've shown that $\{f_n\}_{n=1}^{\infty}$ converges to f in $C_{\mathbb{R}}(X)$ as required.

Now that we know the metric space $C_{\mathbb{R}}(X)$ is complete we can apply the Contraction Lemma 12 to $C_{\mathbb{R}}(X)$:

LEMMA 23. *Suppose that $T : C_{\mathbb{R}}(X) \to C_{\mathbb{R}}(X)$ is a strict contraction on $C_{\mathbb{R}}(X)$, i.e. there is $0 < r < 1$ such that*

$$d(Tf, Tg) \leq r\, d(f, g), \qquad \textit{for all } f, g \in C_{\mathbb{R}}(X).$$

Then T has a unique fixed point h in $C_{\mathbb{R}}(X)$, i.e. there is a unique $h \in C_{\mathbb{R}}(X)$ such that $Th = h$.

2.1. Existence and uniqueness of solutions to initial value problems.

We can use Lemma 23 in the case X is a closed bounded interval in $\mathbb{R}$ to give an elegant proof of a standard existence and uniqueness theorem for solutions to the nonlinear first order initial value problem

$$\left\{ \begin{array}{ccc} y' & = & f(x, y) \\ y(x_0) & = & y_0 \end{array} \right. , \qquad a \leq x \leq b, \tag{2.1}$$

where $a < x_0 < b$, $y_0 \in \mathbb{R}$ and $f : [a, b] \times \mathbb{R} \to \mathbb{R}$ is continuous and satisfies a *Lipschitz condition* in the second variable. A function $h : [\alpha, \beta] \times \mathbb{R} \to \mathbb{R}$ is said to satisfy a *Lipschitz condition* in the second variable if there is a positive number L such that

$$|f(x,y) - f(x,y')| \leq L\,|y-y'|\,, \qquad \text{for all } x \in [\alpha,\beta] \text{ and } y,y' \in \mathbb{R}. \tag{2.2}$$

DEFINITION 35. *A differentiable function $y : [a,b] \to \mathbb{R}$ is defined to be a* solution *to (2.1) if*

$$\begin{aligned} y'(x) &= f(x,y(x)) \text{ for all } x \in [a,b]\,, \\ \text{and } y(x_0) &= y_0. \end{aligned} \tag{2.3}$$

THEOREM 49. *Suppose that $\alpha < x_0 < \beta$, $y_0 \in \mathbb{R}$ and $f : [\alpha,\beta] \times \mathbb{R} \to \mathbb{R}$ is continuous and satisfies the Lipschitz condition (2.2). Then there are a,b satisfying $\alpha \leq a < x_0 < b \leq \beta$ such that there is a unique solution $y : [a,b] \to \mathbb{R}$ to the intial value problem (2.1).*

Proof: Our strategy is to first use the Fundamental Theorem of Calculus to replace the initial value problem (2.1) with an equivalent integral equation (2.4). Then we observe that a solution to the integral equation (2.4) is a *fixed point* of a certain map $T : C_{\mathbb{R}}([\alpha,\beta]) \to C_{\mathbb{R}}([\alpha,\beta])$. Then we will choose $a < x_0 < b$ sufficiently close to x_0 that the map T is a *strict contraction* when viewed as a map on the metric space $C_{\mathbb{R}}([a,b])$. The existence of a unique fixed point to the integral equation (2.4) then follows immediately from Lemma 23, and this proves the theorem. Here are the details.

We claim that $y : [a,b] \to \mathbb{R}$ is differentiable and a solution to (2.1) if and only if y is continuous and satisfies the integral equation

$$y(x) = y_0 + \int_{x_0}^{x} f(t,y(t))\,dt, \qquad a \leq x \leq b. \tag{2.4}$$

This equivalence will use only the continuity of f and not the Lipschitz condition (2.2). Note that if y is continuous, then the map $t \to (t,y(t)) \in \mathbb{R}^2$ is continuous, and hence so is the map $t \to f(t,y(t)) \in \mathbb{R}$. Theorem 41 thus shows that the integrals on the right side of (2.4) all exist when y is continuous.

Suppose first that $y : [a,b] \to \mathbb{R}$ is a solution to (2.1). This means that $y'(t)$ exists on $[a,b]$ and satisfies (2.3). However, $y(t)$ is then also continuous and hence so is $f(t,y(t))$ by the above comments. Thus (2.3) shows that y' is actually continuous on $[a,b]$, hence $y' \in \mathcal{R}[x_0,x]$ for all $a \leq x \leq b$. Now apply the Fundamental Theorem of Calculus 2.1 to (2.3) to obtain

$$y(x) - y(x_0) = \int_{x_0}^{x} y'(t)\,dt = \int_{x_0}^{x} f(t,y(t))\,dt,$$

which is (2.4) since $y(x_0) = y_0$ by the second line in (2.3).

Conversely, suppose that $y : [a,b] \to \mathbb{R}$ is a continuous solution to (2.4). Then the integrand $f(t,y(t))$ is continuous and by Theorem 44 we have

$$y'(x) = \frac{d}{dx}\int_{x_0}^{x} f(t,y(t))\,dt = f(x,y(x))\,, \qquad a \leq x \leq b,$$

which is the first line in (2.3). The second line in (2.3) is immediate upon setting $x = x_0$ in (2.4).

Now we observe that y is a solution to the integral equation (2.4) if and only if $y \in C_{\mathbb{R}}([a,b])$ is a *fixed point* of the map $T : C_{\mathbb{R}}([a,b]) \to C_{\mathbb{R}}([a,b])$ defined by

$$T\varphi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt, \qquad a \leq x \leq b,\ \varphi \in C_{\mathbb{R}}([a,b]).$$

Note that T maps $C_{\mathbb{R}}([a,b])$ to itself since if $\varphi \in C_{\mathbb{R}}([a,b])$ then $f(t,\varphi(t))$ is continuous on $[a,b]$ and Theorems 41 and 44 show that $T\varphi \in C_{\mathbb{R}}([a,b])$. In order to apply Lemma 23 we will need to choose $a < x_0 < b$ sufficiently close to x_0 that the map T is a strict contraction on $C_{\mathbb{R}}([a,b])$. We begin by estimating the distance in $C_{\mathbb{R}}([a,b])$ between $T\varphi$ and $T\psi$ for any pair $\varphi, \psi \in C_{\mathbb{R}}([a,b])$:

$$\begin{aligned} d_{C_{\mathbb{R}}([a,b])}(T\varphi, T\psi) &= \sup_{a\leq x\leq b} |T\varphi(x) - T\psi(x)| \\ &= \sup_{a\leq x\leq b} \left| \int_{x_0}^{x} f(t,\varphi(t)) - f(t,\psi(t))\, dt \right| \\ &\leq \sup_{a\leq x\leq b} \left| \int_{x_0}^{x} |f(t,\varphi(t)) - f(t,\psi(t))|\, dt \right| \\ &\leq \sup_{a\leq x\leq b} \left| \int_{x_0}^{x} L\,|\varphi(t) - \psi(t)|\, dt \right|, \end{aligned}$$

where the final line uses the Lipschitz condition (2.2). But with

$$m \equiv \max\{b - x_0, x_0 - a\},$$

we can dominate the final expression by

$$L \sup_{a\leq x\leq b} \left| \int_{x_0}^{x} |\varphi(t) - \psi(t)|\, dt \right| \leq Lm \sup_{a\leq t\leq b} |\varphi(t) - \psi(t)| = Lm\; d_{C_{\mathbb{R}}([a,b])}(\varphi,\psi).$$

Thus if we choose a and b so close to x_0 that $m < \frac{1}{L}$, then $r \equiv Lm < 1$ and we have

$$d_{C_{\mathbb{R}}([a,b])}(T\varphi, T\psi) \leq r\; d_{C_{\mathbb{R}}([a,b])}(\varphi,\psi),$$

for all $\varphi, \psi \in C_{\mathbb{R}}([a,b])$, which shows that T is a contraction on $C_{\mathbb{R}}([a,b])$ since $r < 1$. Lemma 23 now shows that T has a unique fixed point $y \in C_{\mathbb{R}}([a,b])$, and by what we proved above, this function y is the unique solution to the initial value problem (2.1).

2.1.1. *An example.* Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by $f(x,y) = y$. Then f is continuous and satisfies the Lipschitz condition (2.2) with $L = 1$. Theorem 49 then yields a unique solution $E : [a,b] \to \mathbb{R}$ to the initial value problem

$$\left\{ \begin{array}{rcl} y' &=& y \\ y(0) &=& 1 \end{array} \right., \qquad a \leq x \leq b,$$

for some $a < 0 < b$. An examination of the proof of Theorem 49 shows that we only need a and b to satisfy $m \equiv \max\{b - 0, 0 - a\} < \frac{1}{L} = 1$, so that we have a unique solution $E_\lambda : [-\lambda, \lambda] \to \mathbb{R}$ for any $0 < \lambda < 1$. By uniqueness, all of these solutions E_λ coincide on common intervals of definition. Thus we have a function $E : (-1,1) \to R$ satisfying

$$\left\{ \begin{array}{rcl} E' &=& E \\ E(0) &=& 1 \end{array} \right., \qquad -1 < x < 1.$$

But much more is true. If $-1 < x_0 < 1$ and $0 < \lambda < 1$ then the above reasoning shows that the initial value problem

$$\left\{ \begin{array}{rcl} y' & = & y \\ y(x_0) & = & E(x_0) \end{array} \right. , \qquad -\lambda \leq x - x_0 \leq \lambda,$$

has a unique solution $F_\lambda : [x_0 - \lambda \leq x \leq x_0 + \lambda] \to \mathbb{R}$. But $F_\lambda(x_0) = E(x_0)$ and so by uniqueness we must have $F_\lambda = E$ on their common interval of definition. Repeating this type of argument it follows that there is a unique extension of E to a function E defined on all of $\mathbb{R}$ that satisfies

$$\begin{array}{rcl} E'(x) & = & E(x), \qquad x \in \mathbb{R}, \\ E(0) & = & 1. \end{array}$$

Thus E is infinitely differentiable $E^{(n)} = E$ and is of course the exponential function $Exp(x)$ in (3.1), as can be easily seen using Taylor's formula Theorem 37:

$$\begin{array}{rcl} E(x) & = & E(0) + E'(0)x + \ldots + E^{(n)}(0)\dfrac{x^n}{n!} + E^{(n+1)}(c)\dfrac{x^{n+1}}{(n+1)!} \\ & = & 1 + x + \ldots + \dfrac{x^n}{n!} + E^{(n+1)}(c)\dfrac{x^{n+1}}{(n+1)!}, \end{array}$$

for some c between 0 and x. Indeed,

$$\left| E(c)\frac{x^{n+1}}{(n+1)!} \right| \leq \sup_{|c| \leq |x|} |E(c)| \frac{|x|^{n+1}}{(n+1)!} \to 0$$

as $n \to \infty$, so that

$$E(x) = 1 + x + \ldots + \frac{x^n}{n!} + \ldots = \sum_{n=0}^{\infty} \frac{x^n}{n!} = Exp(x).$$

REMARK 18. *In most applications it is not the case that $f : [\alpha, \beta] \times \mathbb{R} \to \mathbb{R}$ satisfies a Lipschitz condition for all $y, y' \in \mathbb{R}$ as in (2.2), but more likely that the Lipschitz condition is restricted to a finite interval $y, y' \in [\gamma, \delta]$, or even that f is only defined on a bounded rectangle $[\alpha, \beta] \times [\gamma, \delta]$ with $y_0 \in (\gamma, \delta)$. Theorem 49 can still be profitably applied however if we simply redefine $f(x, y)$ to be constant in y outside an interval $[\gamma, \delta]$ that contains y_0 in its interior. More precisely, set*

$$\widetilde{f}(x,y) \equiv \left\{ \begin{array}{lll} f(x,\gamma) & \text{if} \quad \alpha \leq x \leq \beta, & y \leq \gamma, \\ f(x,y) & \text{if} \quad \alpha \leq x \leq \beta, & \gamma \leq y \leq \delta, \\ f(x,\delta) & \text{if} \quad \alpha \leq x \leq \beta, & \delta \leq y, \end{array} \right. .$$

Then if $f : [\alpha, \beta] \times [\gamma, \delta] \to \mathbb{R}$ is continuous and satisfies the local *Lipschitz condition*

$$|f(x,y) - f(x,y')| \leq L|y - y'|, \qquad \text{for all } x \in [\alpha, \beta] \text{ and } y, y' \in [\gamma, \delta],$$

the function $\widetilde{f} : [\alpha, \beta] \times \mathbb{R} \to \mathbb{R}$ is continuous and satisfies the Lipschitz condition (2.2). Thus Theorem 49 produces $a < x_0 < b$ and a solution $y : [a, b] \times \mathbb{R} \to \mathbb{R}$ to the initial value problem

$$\left\{ \begin{array}{rcl} y' & = & \widetilde{f}(x,y) \\ y(x_0) & = & y_0 \end{array} \right. , \qquad a \leq x \leq b.$$

Since y is continuous and

$$(x_0, y(x_0)) = (x_0, y_0) \in (a, b) \times (\gamma, \delta),$$

there exist $a \leq A < x_0 < B \leq b$ *such that* $(x, y(x)) \in [A, B] \times [\gamma, \delta]$ *for* $A \leq x \leq B$. *But then* $\widetilde{f}(x, y(x)) = f(x, y(x))$ *for such* x *and we see from (2.3) that the restriction of* y *to* $[A, B]$ *solves the initial value problem*

$$\left\{ \begin{array}{rcl} y' & = & f(x, y) \\ y(x_0) & = & y_0 \end{array} \right. , \qquad A \leq x \leq B.$$

2.2. Space-filling curves and snowflake curves. We first use the completeness of $C_{\mathbb{R}}([0,1])$ to construct two continuous maps $\varphi, \psi \in C_{\mathbb{R}}([0,1])$ with the property that

$$\{(\varphi(t), \psi(t)) : 0 \leq t \leq 1\} = [0,1] \times [0,1].$$

Thus if we define $\Phi(t) = (\varphi(t), \psi(t))$ for $0 \leq t \leq 1$, then $\Phi : [0,1] \to [0,1]^2$ takes the closed unit interval continuously *onto* the closed unit square! This is the simplest example of a *space-filling* curve. Note that it is impossible for a space-filling curve to be one-to-one:

LEMMA 24. *If* $\Phi : [0,1] \to [0,1]^2$ *is both continuous and onto, then* Φ *is* not *one-to-one.*

Proof: Suppose in order to derive a contradiction that Φ is continuous, one-to-one and onto. Since $[0,1]$ is compact, Corollary 11 then shows that the inverse map $\Phi^{-1} : [0,1]^2 \to [0,1]$ is continuous. Now consider the distinct points $P = \Phi(0)$ and $Q = \Phi(1)$ in the unit square. Pick any two continuous curves $\boldsymbol{\gamma}_j(t) : [0,1] \to [0,1]^2$, $j = 1, 2$, for which

$$\begin{array}{rcl} \boldsymbol{\gamma}_1(0) & = & \boldsymbol{\gamma}_2(1) = P, \\ \boldsymbol{\gamma}_1(1) & = & \boldsymbol{\gamma}_2(0) = Q, \\ \boldsymbol{\gamma}_1(t) & \neq & \boldsymbol{\gamma}_2(t), \qquad 0 \leq t \leq 1. \end{array} \tag{2.5}$$

Thus $\boldsymbol{\gamma}_1$ takes P to Q continuously and $\boldsymbol{\gamma}_2$ takes Q to P continuously, and the images $\boldsymbol{\gamma}_1(t)$ and $\boldsymbol{\gamma}_2(t)$ of the two curves in the square are distinct for each t.

Now consider the difference of the composition of these two curves with the continuous map Φ^{-1}:

$$\beta(t) = \Phi^{-1}(\boldsymbol{\gamma}_1(t)) - \Phi^{-1}(\boldsymbol{\gamma}_2(t)), \qquad 0 \leq t \leq 1.$$

Thus $\beta : [0,1] \to [0,1]$ is continuous and

$$\begin{array}{rcl} \beta(0) & = & \Phi^{-1}(P) - \Phi^{-1}(Q) = -1, \\ \beta(1) & = & \Phi^{-1}(Q) - \Phi^{-1}(P) = 1. \end{array}$$

Since 0 is an intermediate value, the Intermediate Value Theorem shows that there is $c \in (0,1)$ such that

$$0 = \beta(c) = \Phi^{-1}(\boldsymbol{\gamma}_1(c)) - \Phi^{-1}(\boldsymbol{\gamma}_2(c)),$$

which implies

$$\boldsymbol{\gamma}_1(c) = \Phi\left(\Phi^{-1}(\boldsymbol{\gamma}_1(c))\right) = \Phi\left(\Phi^{-1}(\boldsymbol{\gamma}_2(c))\right) = \boldsymbol{\gamma}_2(c),$$

contradicting the third line in (2.5).

To construct our space-filling curve $\Phi(t) = (\varphi(t), \psi(t))$, we begin with a continuous function $f : \mathbb{R} \to [0,1]$ of period 2, i.e. $f(t+2) = f(t)$ for all $t \in \mathbb{R}$, that satisfies

$$f(t) = \left\{ \begin{array}{lll} 0 & \text{if} & 0 \leq t \leq \frac{1}{3} \\ 1 & \text{if} & \frac{2}{3} \leq t \leq 1 \end{array} \right. .$$

Then for $N \in \mathbb{N}$ define

$$\varphi_N(t) = \sum_{n=1}^{N} \frac{1}{2^n} f\left(3^{2n-1}t\right) \text{ and } \psi_N(t) = \sum_{n=1}^{N} \frac{1}{2^n} f\left(3^{2n}t\right), \qquad 0 \le t \le 1.$$

Each of the sequences $\{\varphi_N\}_{N=1}^{\infty}$ and $\{\psi_N\}_{N=1}^{\infty}$ is Cauchy in the metric space $C_{\mathbb{R}}([0,1])$ since if $M < N$,

$$\begin{aligned} d(\varphi_M, \varphi_N) &= \sup_{0\le t\le 1} |\varphi_M(t) - \varphi_N(t)| \\ &= \sup_{0\le t\le 1} \left| \sum_{n=M+1}^{N} \frac{1}{2^n} f\left(3^{2n-1}t\right) \right| \le \sum_{n=M+1}^{N} \frac{1}{2^n} < \frac{1}{2^M} \end{aligned}$$

tends to 0 as $M \to \infty$, and similarly $d(\psi_M, \psi_N) \to 0$ as $M \to \infty$. Since $C_{\mathbb{R}}([0,1])$ is complete, there are continuous functions φ and ψ on $[0,1]$ such that

$$\varphi = \lim_{N\to\infty} \varphi_N \text{ and } \psi = \lim_{N\to\infty} \psi_N \text{ in } C_{\mathbb{R}}([0,1]).$$

Then $\Phi(t) = (\varphi(t), \psi(t))$, $0 \le t \le 1$, defines a continuous map from $[0,1]$ into the unit square $[0,1]^2$ since $0 \le \varphi(t), \psi(t) \le 1$ for $0 \le t \le 1$. We claim that given $(x_0, y_0) \in [0,1]^2$ there is $t_0 \in [0,1]$ such that $\Phi(t_0) = (x_0, y_0)$. To see this expand both x_0 and y_0 in binary series:

$$x_0 = \sum_{n=1}^{\infty} a_{2n-1} \left(\frac{1}{2}\right)^n \text{ and } y_0 = \sum_{n=1}^{\infty} a_{2n} \left(\frac{1}{2}\right)^n,$$

where each coefficient a_{2n-1} and a_{2n} is either 0 or 1. Now set

$$t_0 = \sum_{k=1}^{\infty} \frac{2}{3^{k+1}} a_k.$$

For $\ell \in \mathbb{N}$ consider the number

$$3^\ell t_0 = \sum_{k=1}^{\ell-1} 3^{\ell-k-1} 2a_k + \frac{2}{3} a_\ell + \sum_{k=\ell+1}^{\infty} 3^{\ell-k-1} 2a_k = A_\ell + \frac{2}{3} a_\ell + B_\ell.$$

Now $A_\ell = \sum_{k=1}^{\ell-1} 3^{\ell-k-1} 2a_k$ is an even integer and

$$\begin{aligned} 0 &\le B_\ell = \sum_{k=\ell+1}^{\infty} 3^{\ell-k-1} 2a_k \le \sum_{k=\ell+1}^{\infty} 3^{\ell-k-1} 2 \\ &= 2\left\{\frac{1}{9} + \frac{1}{27} + \ldots\right\} = 2\frac{1}{9}\frac{1}{1-\frac{1}{3}} = \frac{1}{3}. \end{aligned}$$

If follows from the fact that f has period 2 that

$$f\left(3^\ell t_0\right) = f\left(A_\ell + \frac{2}{3}a_\ell + B_\ell\right) = f\left(\frac{2}{3}a_\ell + B_\ell\right),$$

and then from the fact that f is constant on $\left[0, \frac{1}{3}\right]$ and $\left[\frac{2}{3}, 1\right]$ that

$$f\left(3^\ell t_0\right) = f\left(\frac{2}{3}a_\ell\right),$$

and finally that

$$f\left(3^\ell t_0\right) = a_\ell, \tag{2.6}$$

since $f(0) = 0$ and $f\left(\frac{2}{3}\right) = 1$.

Armed with (2.6) we obtain

$$\begin{aligned}
\varphi(t_0) &= \lim_{N\to\infty} \varphi_N(t_0) = \sum_{n=1}^{\infty} \frac{1}{2^n} f\left(3^{2n-1}t_0\right) = \sum_{n=1}^{\infty} \frac{1}{2^n} a_{2n-1} = x_0, \\
\psi(t_0) &= \lim_{N\to\infty} \psi_N(t_0) = \sum_{n=1}^{\infty} \frac{1}{2^n} f\left(3^{2n}t_0\right) = \sum_{n=1}^{\infty} \frac{1}{2^n} a_{2n} = y_0,
\end{aligned}$$

which implies $\Phi(t_0) = (\varphi(t_0), \psi(t_0)) = (x_0, y_0)$, and completes the proof that Φ maps $[0,1]$ onto $[0,1]^2$.

Now we return to the von Koch snowflake K constructed in Subsection 3.2 of Chapter 3. Recall that we constructed the snowflake in a sequence of steps that we called 'generations'. At the k^{th} generation, we had constructed a polygonal path consisting of 4^k closed segments$\left\{L_j^k\right\}_{j=1}^{4^k}$ each of length $\frac{1}{3^k}$. We denoted this polygonal 'snowflake-shaped' path by P_k. We now parameterize this polygonal path P_k with a *constant speed* parameterization on the unit interval $[0,1]$. Since the length of P_k is

$$\mathit{length}\ \ (P_k) = 4^k \cdot \frac{1}{3^k} = \left(\frac{4}{3}\right)^k,$$

this will result in a curve

$$\boldsymbol{\gamma}_k(t) = (\alpha_k(t), \beta_k(t)), \qquad 0 \le t \le 1,$$

that traces out the polygonal path P_k in such a way that

$$\left\|\left(\alpha_k'(t), \beta_k'(t)\right)\right\| = \sqrt{\left|\alpha_k'(t)\right|^2 + \left|\beta_k'(t)\right|^2} = \left(\frac{4}{3}\right)^k,$$

at all t except those corresponding to the vertices of P_k.

We now observe that the vertices of P_k are precisely the points $\boldsymbol{\gamma}_k\left(\frac{j}{4^k}\right)$, and moreover that

$$\boldsymbol{\gamma}_{k'}\left(\frac{j}{4^k}\right) = \boldsymbol{\gamma}_k\left(\frac{j}{4^k}\right) \text{ whenever } k' \ge k.$$

Thus the vertices in the constructions remain fixed once they appear, and are thereafter achieved by each $\boldsymbol{\gamma}_{k'}$ with the same parameter value. In fact we can prove the following estimate for the difference between consecutive curves by induction:

$$\left|\boldsymbol{\gamma}_{k+1}(t) - \boldsymbol{\gamma}_k(t)\right| \le \frac{1}{3^k}, \qquad 0 \le t \le 1,\ k \ge 1.$$

As a consequence we see that each of the sequences $\{\alpha_k\}_{k=1}^{\infty}$ and $\{\beta_k\}_{k=1}^{\infty}$ of continuous functions on $[0,1]$ is a Cauchy sequence in the metric space $C_{\mathbb{R}}([0,1])$. Indeed, if $m < n$ then the triangle inequality gives

$$\begin{aligned}
d(\alpha_m, \alpha_n) &\le \sum_{k=m}^{n-1} d(\alpha_k, \alpha_{k+1}) \le \sum_{k=m}^{n-1} \sup_{0\le t\le 1} \left|\alpha_{k+1}(t) - \alpha_k(t)\right| \\
&\le \sum_{k=m}^{n-1} \sup_{0\le t\le 1} \left|\boldsymbol{\gamma}_{k+1}(t) - \boldsymbol{\gamma}_k(t)\right| \le \sum_{k=m}^{\infty} \frac{1}{3^k} = \frac{1}{3^m} \frac{1}{1-\frac{1}{3}},
\end{aligned}$$

which tends to 0 as $m \to \infty$, and similarly for $d(\beta_m, \beta_n)$. Thus there are continuous functions $\alpha, \beta \in C_{\mathbb{R}}([0,1])$ such that the curve

$$\boldsymbol{\gamma}(t) = \lim_{k\to\infty} \boldsymbol{\gamma}_k(t) = \lim_{k\to\infty} (\alpha_k(t), \beta_k(t)) = \left(\lim_{k\to\infty} \alpha_k(t), \lim_{k\to\infty} \beta_k(t)\right) = (\alpha(t), \beta(t))$$

maps *onto* the von Koch snowflake K.

We now sketch a proof that $\boldsymbol{\gamma} : [0,1] \to K$ is one-to-one, thus demonstrating that the fractal K is a closed *Jordan arc*, namely a continuous one-to-one image of the closed unit interval $[0,1]$. Indeed, let S_1, S_2, S_3 and S_4 be the similarities characterizing K in Theorem 13. These are given in the table in Subsection 3.2 of Chapter 3: for $x = (x_1, x_2) \in \mathbb{R}^2$,

$$\begin{aligned} S_1 x &= \frac{1}{3} x, \\ S_2 x &= \frac{1}{3}(M_2 x + (1,0)), \\ S_3 x &= \frac{1}{3}\left(M_3 x + \left(\frac{3}{2}, \frac{\sqrt{3}}{2}\right)\right), \\ S_4 x &= \frac{1}{3}(x + (2,0)). \end{aligned}$$

Now define T to be the *open* triangle with vertices

$$(0,0), \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), (1,0),$$

and for $0 \le t \le 1$, expand t in a series

$$t = \sum_{n=1}^{\infty} a_n \left(\frac{1}{4}\right)^n, \qquad a_n \in \{0,1,2,3\}, \tag{2.7}$$

where the sequence $\{a_n\}_{n=1}^{\infty}$ does *not* end in an infinite string of consecutive 3's, except for the case where *all* the a_n are 3. With this restriction, the series representation (2.7) of $t \in [0,1]$ is unique.

One can now show (we leave this to the reader) that the intersection

$$\begin{aligned} &\bigcap_{j=1}^{\infty} S_{a_j+1}\left(...S_{a_2+1}\left(S_{a_1+1}\left(\overline{T}\right)\right)\right) \\ = \quad &S_{a_1+1}\left(\overline{T}\right) \cap S_{a_2+1}\left(S_{a_1+1}\left(\overline{T}\right)\right) \cap ... \cap S_{a_j+1}\left(...S_{a_2+1}\left(S_{a_1+1}\left(\overline{T}\right)\right)\right) \cap ... \end{aligned}$$

consists of *exactly* the single point $\boldsymbol{\gamma}(t)$. Moreover:

- The four triangles $S_1(T)$, $S_2(T)$, $S_3(T)$, $S_4(T)$ are pairwise disjoint, as well as the four triangles

  $$S_1(S(T)), \ S_2(S(T)), \ S_3(S(T)), \ S_4(S(T))$$

  where S is any finite composition of the similarities S_1, S_2, S_3 and S_4.

It now follows easily that $\boldsymbol{\gamma}(t) \neq \boldsymbol{\gamma}(t')$ for $t \neq t'$ upon expanding

$$t = \sum_{n=1}^{\infty} a_n \left(\frac{1}{4}\right)^n \text{ and } t' = \sum_{n=1}^{\infty} a'_n \left(\frac{1}{4}\right)^n$$

as in (2.7) above, considering the smallest n for which $a_n \neq a'_n$, and then applying the observation in the bullet item to $S_{a_n}(S(T))$ and $S_{a'_n}(S(T))$ where $S = S_{a_{n-1}+1} \circ ... \circ S_{a_2+1} \circ S_{a_1+1}$. This shows that $S_{a_n}(S(T)) \cap S_{a'_n}(S(T)) = \emptyset$ and in order to obtain $S_{a_n}\left(S\left(\overline{T}\right)\right) \cap S_{a'_n}\left(S\left(\overline{T}\right)\right) = \emptyset$, we use the assumption that the coefficients in the series representation (2.7) do not end in an infinite string of consecutive 3's.

Finally, we show that the curve $\boldsymbol{\gamma}(t)$ is nowhere differentiable. For each k there is j such that

$$\frac{j}{4^k} \leq t < \frac{j+1}{4^k}.$$

Let $A_k = \frac{j}{4^k}$ and $B_k = \frac{j+1}{4^k}$. Suppose in order to derive a contradiction that $\boldsymbol{\gamma}'(t)$ exists. Then we would have

$$\lim_{k\to\infty} \left\| \frac{\boldsymbol{\gamma}(B_k) - \boldsymbol{\gamma}(A_k)}{B_k - A_k} \right\| = \left\| \lim_{k\to\infty} \frac{\boldsymbol{\gamma}(B_k) - \boldsymbol{\gamma}(A_k)}{B_k - A_k} \right\| = \left\| \boldsymbol{\gamma}'(t) \right\|. \tag{2.8}$$

However, the length of the line segment $\boldsymbol{\gamma}(B_k) - \boldsymbol{\gamma}(A_k)$ is $\frac{1}{3^k}$, and $B_k - A_k = \frac{1}{4^k}$, so

$$\left\| \frac{\boldsymbol{\gamma}(B_k) - \boldsymbol{\gamma}(A_k)}{B_k - A_k} \right\| = \frac{\frac{1}{3^k}}{\frac{1}{4^k}} = \left(\frac{4}{3}\right)^k,$$

which tends to ∞ as $k \to \infty$, the desired contradiction.

CHAPTER 8

Lebesgue measure theory

Recall that f is Riemann integrable on $[0,1)$, written $f \in \mathcal{R}[0,1)$, if $\mathcal{U}(f) = \mathcal{L}(f)$, and we denote the common value by $\int_0^1 f$ or $\int_0^1 f(x)\,dx$. Here $\mathcal{U}(f)$ and $\mathcal{L}(f)$ are the upper and lower Riemann integrals of f on $[0,1)$ respectively given by

$$\begin{aligned} \mathcal{U}(f) &= \inf_{\mathcal{P}} \sum_{n=1}^{N} \left(\sup_{[x_{n-1},x_n)} f \right) \triangle x_n, \\ \mathcal{L}(f) &= \sup_{\mathcal{P}} \sum_{n=1}^{N} \left(\inf_{[x_{n-1},x_n)} f \right) \triangle x_n, \end{aligned}$$

where $\mathcal{P} = \{0 = x_0 < x_1 < ... < x_N = 1\}$ is any partition of $[0,1)$ and $\triangle x_n = x_n - x_{n-1} > 0$. For convenience we work with $[0,1)$ in place of $[0,1]$ for now.

This definition is simple and easy to work with and applies in particular to bounded continuous functions f on $[0,1)$ since it is not too hard to prove that $f \in \mathcal{R}[0,1)$ for such f. However, if we consider the vector space $L^2_{\mathcal{R}}([0,1))$ of Riemann integrable functions $f \in \mathcal{R}[0,1)$ endowed with the metric

$$d(f,g) = \left(\int_0^1 |f(x) - g(x)|^2 \, dx \right)^{\frac{1}{2}},$$

it turns out that while $L^2_{\mathcal{R}}([0,1))$ can indeed be proved a metric space, it fails to be *complete*. This is a serious shortfall of Riemann's theory of integration, and is our main motivation for considering the more complicated theory of Lebesgue below. We note that the immediate reason for the lack of completeness of $L^2_{\mathcal{R}}([0,1))$ is the inability of Riemann's theory to handle general unbounded functions. However, even locally there are problems. For example, once we have Lebesgue's theory in hand, we can construct a famous example of a Lebesgue measurable subset E of $[0,1)$ with the (somewhat surprising) property that

$$0 < |E \cap (a,b)| < b - a, \qquad 0 \le a < b \le 1,$$

where $|F|$ denotes the Lebesgue measure of a measurable set F (see Problem 5 below). It follows that the characteristic function χ_E is bounded and Lebesgue measurable, but that there is no Riemann integrable function f such that $f = \chi_E$ almost everywhere, since such an f would satisfy $\mathcal{U}(f) = 1$ and $\mathcal{L}(f) = 0$. Nevertheless, by Lusin's Theorem (see page 34 in [**5**] or page 55 in [**4**]) there is a sequence of compactly supported *continuous* functions (hence Riemann integrable) converging to χ_E almost everywhere.

On the other hand, in *Lebesgue's* theory of integration, we partition the *range* $[0,M)$ of the function into a homogeneous partition,

$$[0,M)=\bigcup_{n=1}^{N}\left[(n-1)\frac{M}{N},n\frac{M}{N}\right)\equiv\bigcup_{n=1}^{N}I_n,$$

and we consider the associated upper and lower *Lebesgue* sums of f on $[0,1)$ defined by

$$\begin{aligned}U^*(f;\mathcal{P}) &= \sum_{n=1}^{N}\left(n\frac{M}{N}\right)\left|f^{-1}(I_n)\right|,\\ L^*(f;\mathcal{P}) &= \sum_{n=1}^{N}\left((n-1)\frac{M}{N}\right)\left|f^{-1}(I_n)\right|,\end{aligned}$$

where of course

$$f^{-1}(I_n)=\left\{x\in[0,1):f(x)\in I_n=\left[(n-1)\frac{M}{N},n\frac{M}{N}\right)\right\},$$

and $|E|$ denotes the "measure" or "length" of the subset E of $[0,1)$.

Here there will be no problem obtaining that $U^*(f;\mathcal{P})-L^*(f;\mathcal{P})$ is small provided we can make sense of $\left|f^{-1}(I_n)\right|$. But this is precisely the difficulty with Lebesgue's approach - we need to define a notion of "measure" or "length" for subsets E of $[0,1)$. That this is not going to be as easy as we might hope is evidenced by the following negative result. Let $\mathcal{P}([0,1))$ denote the power set of $[0,1)$, i.e. the set of all subsets of $[0,1)$. For $x\in[0,1)$ and $E\in\mathcal{P}([0,1))$ we define the translation $E\oplus x$ of E by x to be the set in $\mathcal{P}([0,1))$ defined by

$$\begin{aligned}E\oplus x &= E+x \qquad (\operatorname{mod}1)\\ &= \{z\in[0,1):\text{there is }y\in E\text{ with }y+x-z\in\mathbb{Z}\}.\end{aligned}$$

Theorem 50. *There is* no *map* $\mu:\mathcal{P}([0,1))\to[0,\infty)$ *satisfying the following three properties:*

1. $\mu([0,1))=1$,
2. $\mu\left(\dot{\bigcup}_{n=1}^{\infty}E_n\right)=\sum_{n=1}^{\infty}\mu(E_n)$ *whenever* $\{E_n\}_{n=1}^{\infty}$ *is a pairwise disjoint sequence of sets in* $\mathcal{P}([0,1))$,
3. $\mu(E\oplus x)=\mu(E)$ *for all* $E\in\mathcal{P}([0,1))$.

Remark 19. *All three of these properties are desirable for any notion of measure or length of subsets of* $[0,1)$*. The theorem suggests then that we should* not *demand that every subset of* $[0,1)$ *be "measurable". This will then restrict the functions* f *that we can integrate to those for which* $f^{-1}([a,b))$ *is "measurable" for all* $-\infty<a<b<\infty$.

Proof: Let $\{r_n\}_{n=1}^{\infty}=\mathbb{Q}\cap[0,1)$ be an enumeration of the rational numbers in $[0,1)$. Define an equivalence relation on $[0,1)$ by declaring that $x\sim y$ if $x-y\in\mathbb{Q}$. Let $\mathcal{A}$ be the set of equivalence classes. Use the *axiom of choice* to pick a representative $a=\langle A\rangle$ from each equivalence class A in $\mathcal{A}$. Finally, let $E=\{\langle A\rangle:A\in\mathcal{A}\}$ be the set consisting of these representatives a, one from each equivalence class A in $\mathcal{A}$.

Then we have

$$[0,1) = \dot{\bigcup}_{n=1}^{\infty} E \oplus r_n.$$

Indeed, if $x \in [0,1)$, then $x \in A$ for some $A \in \mathcal{A}$, and thus $x \sim a = \langle A \rangle$, i.e. $x - a \in \{r_n\}_{n=1}^{\infty}$. If $x \geq a$ then $x - a \in \mathbb{Q} \cap [0,1)$ and $x = a + r_m$ where $a \in E$ and $r_m \in \{r_n\}_{n=1}^{\infty}$. If $x < a$ then $x - a + 1 \in \mathbb{Q} \cap [0,1)$ and $x = a + (r_m \ominus 1)$ where $a \in E$ and $r_m \ominus 1 \in \{r_n\}_{n=1}^{\infty}$. Finally, if $a \oplus r_m = b \oplus r_n$, then $a \ominus b = r_n \ominus r_m \in \mathbb{Q}$ which implies that $a \sim b$ and then $r_n = r_m$.

Now by properties (1), (2) and (3) in succession we have

$$1 = \mu([0,1)) = \mu\left(\dot{\bigcup}_{n=1}^{\infty} E \oplus r_n\right) = \sum_{n=1}^{\infty} \mu(E \oplus r_n) = \sum_{n=1}^{\infty} \mu(E),$$

which is impossible since the infinite series $\sum_{n=1}^{\infty} \mu(E)$ is either ∞ if $\mu(E) > 0$ or 0 if $\mu(E) = 0$.

1. Lebesgue measure on the real line

In order to define a "measure" satisfying the three properties in Theorem 50, we must restrict the domain of definition of the set functional μ to a "suitable" proper subset of the power set $\mathcal{P}([0,1))$. A good notion of "suitable" is captured by the following definition where we expand our quest for measure to the entire real line.

DEFINITION 36. *A collection $\mathcal{A} \subset \mathcal{P}(\mathbb{R})$ of subsets of real numbers $\mathbb{R}$ is called a σ-*algebra *if the following properties are satisfied:*

(1) *$\phi \in \mathcal{A}$,*
(2) *$A^c \in \mathcal{A}$ whenever $A \in \mathcal{A}$,*
(3) *$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ whenever $A_n \in \mathcal{A}$ for all n.*

Here is the theorem asserting the existence of "Lebesgue measure" on the real line.

THEOREM 51. *There is a σ-algebra $\mathcal{L} \subset \mathcal{P}(\mathbb{R})$ and a function $\mu : \mathcal{L} \to [0,\infty]$ such that*

(1) *$[a,b) \in \mathcal{L}$ and $\mu([a,b)) = b - a$ for all $-\infty < a < b < \infty$,*
(2) *$\dot{\bigcup}_{n=1}^{\infty} E_n \in \mathcal{L}$ and $\mu\left(\dot{\bigcup}_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$ whenever $\{E_n\}_{n=1}^{\infty}$ is a pairwise disjoint sequence of sets in $\mathcal{L}$,*
(3) *$E + x \in \mathcal{L}$ and $\mu(E + x) = \mu(E)$ for all $E \in \mathcal{L}$,*
(4) *$E \in \mathcal{L}$ and $\mu(E) = 0$ whenever $E \subset F$ and $F \in \mathcal{L}$ with $\mu(F) = 0$.*

It turns out that both the σ-algebra $\mathcal{L}$ and the function μ are uniquely determined by these four properties, but we will only need the *existence* of such $\mathcal{L}$ and μ. The sets in the σ-algebra $\mathcal{L}$ are called *Lebesgue measurable* sets.

A pair $(\mathcal{L}, \mu)$ satisfying only property (2) is called a measure space. Property (1) says that the measure μ is an extension of the usual length function on intervals. Property (3) says that the measure is translation invariant, while property (4) says that the measure is complete.

From property (2) and the fact that μ is nonnegative, we easily obtain the following elementary consequences (where membership in $\mathcal{L}$ is implied by *context*):

$$\begin{aligned} \phi &\in \mathcal{L} \text{ and } \mu(\phi) = 0, \\ E &\in \mathcal{L} \text{ for every open set } E \text{ in } \mathbb{R}, \\ \mu(I) &= b - a \text{ for } \textit{any} \text{ interval } I \text{ with endpoints } a \text{ and } b, \\ \mu(E) &= \sup_n \mu(E_n) = \lim_{n\to\infty} \mu(E_n) \text{ if } E_n \nearrow E, \\ \mu(E) &= \inf_n \mu(E_n) = \lim_{n\to\infty} \mu(E_n) \text{ if } E_n \searrow E \textbf{ and } \mu(E_1) < \infty. \end{aligned} \tag{1.1}$$

For example, the fourth line follows from writing

$$E = E_1 \dot{\cup} \left\{ \dot{\bigcup}_{n=1}^{\infty} E_{n+1} \cap (E_n)^c \right\}$$

and then using property (2) of μ.

To prove Theorem 51 we follow the treatment in [**5**] with simplifications due to the fact that Theorem 31 implies the connected open subsets of the real numbers $\mathbb{R}$ are just the open intervals (a, b). Define for *any* $E \in \mathcal{P}(\mathbb{R})$, the *outer Lebesgue measure* $\mu^*(E)$ of E by,

$$\mu^*(E) = \inf \left\{ \sum_{n=1}^{\infty} (b_n - a_n) : E \subset \dot{\bigcup}_{n=1}^{\infty} (a_n, b_n) \text{ and } -\infty \le a_n < b_n \le \infty \right\}.$$

It is immediate that μ^* is monotone,

$$\mu^*(E) \le \mu^*(F) \text{ if } E \subset F.$$

A little less obvious is countable subadditivity of μ^*.

Lemma 25. *μ^* is countably subadditive:*

$$\mu^* \left(\bigcup_{n=1}^{\infty} E_n \right) \le \sum_{n=1}^{\infty} \mu^*(E_n), \qquad \{E_n\}_{n=1}^{\infty} \subset \mathcal{P}(\mathbb{R}).$$

Proof: Given $0 < \varepsilon < 1$, we have $E_n \subset \dot{\bigcup}_{k=1}^{\infty} (a_{k,n}, b_{k,n})$ with

$$\sum_{k=1}^{\infty} (b_{k,n} - a_{k,n}) < \mu^*(E_n) + \frac{\varepsilon}{2^n}, \qquad n \ge 1.$$

Now let

$$\bigcup_{n=1}^{\infty} \left(\dot{\bigcup}_{k=1}^{\infty} (a_{k,n}, b_{k,n}) \right) = \dot{\bigcup}_{m=1}^{M^*} (c_m, d_m),$$

where $M^* \in \mathbb{N} \cup \{\infty\}$. Then define disjoint sets of indices

$$\mathcal{I}_m = \{(k, n) : (a_{k,n}, b_{k,n}) \subset (c_m, d_m)\}.$$

In the case $c_m, d_m \in \mathbb{R}$, we can choose by compactness a finite subset $\mathcal{F}_m$ of $\mathcal{I}_m$ such that

$$\left[c_m + \frac{\varepsilon}{2} \delta_m, d_m - \frac{\varepsilon}{2} \delta_m \right] \subset \bigcup_{(k,n) \in \mathcal{F}_m}^{\infty} (a_{k,n}, b_{k,n}), \tag{1.2}$$

where $\delta_m = d_m - c_m$. Fix m and arrange the left endpoints $\{a_{k,n}\}_{(k,n) \in \mathcal{F}_m}$ in strictly increasing order $\{a_i\}_{i=1}^{I}$ and denote the corresponding right endpoints by

b_i (if there is more than one interval (a_i, b_i) with the same left endpoint a_i, discard all but one of the largest of them). From (1.2) it now follows that $a_{i+1} \in (a_i, b_i)$ for $i < I$ since otherwise b_i would be in the left side of (1.2), but not in the right side, a contradiction. Thus $a_{i+1} - a_i \leq b_i - a_i$ for $1 \leq i < I$ and we have the inequality

$$\begin{aligned}
(1-\varepsilon)\,\delta_m &= \left(d_m - \frac{\varepsilon}{2}\delta_m\right) - \left(c_m + \frac{\varepsilon}{2}\delta_m\right) \\
&\leq b_I - a_1 = (b_I - a_I) + \sum_{i=1}^{I-1}(a_{i+1} - a_i) \\
&\leq \sum_{i=1}^{I}(b_i - a_i) = \sum_{(k,n)\in\mathcal{F}_m}(b_{k,n} - a_{k,n}) \\
&\leq \sum_{(k,n)\in\mathcal{I}_m}(b_{k,n} - a_{k,n}).
\end{aligned}$$

We also observe that a similar argument shows that $\sum_{(k,n)\in\mathcal{I}_m}(b_{k,n} - a_{k,n}) = \infty$ if $\delta_m = \infty$. Then we have

$$\begin{aligned}
\mu^*(E) &\leq \sum_{m=1}^{\infty}\delta_m \leq \frac{1}{1-\varepsilon}\sum_{m=1}^{\infty}\sum_{(k,n)\in\mathcal{F}_m}(b_{k,n} - a_{k,n}) \\
&\leq \frac{1}{1-\varepsilon}\sum_{k,n}(b_{k,n} - a_{k,n}) = \frac{1}{1-\varepsilon}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}(b_{k,n} - a_{k,n}) \\
&< \frac{1}{1-\varepsilon}\sum_{n=1}^{\infty}\left(\mu^*(E_n) + \frac{\varepsilon}{2^n}\right) = \frac{1}{1-\varepsilon}\sum_{n=1}^{\infty}\mu^*(E_n) + \frac{\varepsilon}{1-\varepsilon}.
\end{aligned}$$

Let $\varepsilon \to 0$ to obtain the countable subadditivity of μ^*.

Now define the subset $\mathcal{L}$ of $\mathcal{P}(\mathbb{R})$ to consist of all subsets A of the real line such that for every $\varepsilon > 0$, there is an open set $G \supset A$ satisfying

$$\mu^*(G \setminus A) < \varepsilon. \tag{1.3}$$

Remark 20. *Condition (1.3) says that A can be well approximated from the outside by open sets. The most difficult task we will face below in using this definition of $\mathcal{L}$ is to prove that such sets A can also be well approximated from the inside by closed sets.*

Set

$$\mu(A) = \mu^*(A), \qquad A \in \mathcal{L}.$$

Trivially, every open set and every interval is in $\mathcal{L}$. We will use the following two claims in the proof of Theorem 51.

Claim 2. *If G is open and $G = \dot{\bigcup}_{n=1}^{N^*}(a_n, b_n)$ (where $N^* \in \mathbb{N} \cup \{\infty\}$) is the decomposition of G into its connected components (a_n, b_n) (Proposition 14 of Chapter 5), then*

$$\mu(G) = \mu^*(G) = \sum_{n=1}^{N^*}(b_n - a_n).$$

We first prove Claim 2 when $N^* < \infty$. If $G \subset \dot{\bigcup}_{m=1}^{\infty} (c_m, d_m)$, then for each $1 \leq n \leq N^*$, $(a_n, b_n) \subset (c_m, d_m)$ for some m since (a_n, b_n) is connected. If

$$\mathcal{I}_m = \{n : (a_n, b_n) \subset (c_m, d_m)\},$$

it follows upon arranging the a_n in increasing order that

$$\sum_{n \in \mathcal{I}_m} (b_n - a_n) \leq d_m - c_m,$$

since the intervals (a_n, b_n) are pairwise disjoint. We now conclude that

$$\begin{aligned} \mu^* (G) &= \inf \left\{ \sum_{m=1}^{\infty} (d_m - c_m) : G \subset \dot{\bigcup}_{m=1}^{\infty} (c_m, d_m) \right\} \\ &\geq \sum_{m=1}^{\infty} \sum_{n \in \mathcal{I}_m} (b_n - a_n) = \sum_{n=1}^{N^*} (b_n - a_n), \end{aligned}$$

and hence that $\mu^* (G) = \sum_{n=1}^{N^*} (b_n - a_n)$ by definition since $G \subset \dot{\bigcup}_{m=1}^{N^*} (a_n, b_n)$.

Finally, if $N^* = \infty$, then from what we just proved and monotonicity, we have

$$\mu^* (G) \geq \mu^* \left(\dot{\bigcup}_{m=1}^{N} (a_n, b_n) \right) = \sum_{n=1}^{N} (b_n - a_n)$$

for each $1 \leq N < \infty$. Taking the supremum over N gives $\mu^* (G) \geq \sum_{n=1}^{\infty} (b_n - a_n)$, and then equality follows by definition since $G \subset \dot{\bigcup}_{n=1}^{\infty} (a_n, b_n)$.

Claim 3. *If A and B are disjoint compact subsets of $\mathbb{R}$, then*

$$\mu^* (A) + \mu^* (B) = \mu^* (A \cup B).$$

First note that

$$\delta = dist\,(A, B) \equiv \inf \{|x - y| : x \in A, y \in B\} > 0,$$

since the function $f(x, y) \equiv |x - y|$ is positive and continuous on the closed and bounded (hence compact) subset $A \times B$ of the plane - Theorem 28 shows that f achieves its infimum $dist\,(A, B)$, which is thus positive. So we can find open sets U and V such that

$$A \subset U \text{ and } B \subset V \text{ and } U \cap V = \phi.$$

For example, $U = \bigcup_{x \in A} B\left(x, \frac{\delta}{2}\right)$ and $V = \bigcup_{x \in B} B\left(x, \frac{\delta}{2}\right)$ work. Now suppose that

$$A \cup B \subset G \equiv \dot{\bigcup}_{n=1}^{\infty} (a_n, b_n).$$

Then we have

$$A \subset U \cap G = \dot{\bigcup}_{k=1}^{K^*} (e_k, f_k) \text{ and } B \subset V \cap G = \dot{\bigcup}_{\ell=1}^{L^*} (g_\ell, h_\ell),$$

and then from Claim 2 and monotonicity of μ^* we obtain

$$\begin{aligned}
\mu^*(A)+\mu^*(B) &\leq \sum_{k=1}^{K^*}(f_k-e_k)+\sum_{\ell=1}^{L^*}(h_\ell-g_\ell) \\
&= \mu^*\left(\left(\dot{\bigcup}_{k=1}^{K^*}(e_k,f_k)\right)\dot{\cup}\left(\dot{\bigcup}_{\ell=1}^{L^*}(g_\ell,h_\ell)\right)\right) \\
&\leq \mu^*(G)=\sum_{n=1}^{\infty}(b_n-a_n).
\end{aligned}$$

Taking the infimum over such G gives $\mu^*(A)+\mu^*(B)\leq\mu^*(A\cup B)$, and subadditivity of μ^* now proves equality.

Proof (of Theorem 51): We now prove that $\mathcal{L}$ is a σ-algebra and that $\mathcal{L}$ and μ satisfy the four properties in the statement of Theorem 51. First we establish that $\mathcal{L}$ is a σ-algebra in four steps.

Step 1: $A\in\mathcal{L}$ if $\mu^*(A)=0$.

Given $\varepsilon>0$, there is an open $G\supset A$ with $\mu^*(G)<\varepsilon$. But then $\mu^*(G\setminus A)\leq\mu^*(G)<\varepsilon$ by monontonicity.

Step 2: $\bigcup_{n=1}^{\infty}A_n\in\mathcal{L}$ whenever $A_n\in\mathcal{L}$ for all n.

Given $\varepsilon>0$, there is an open $G_n\supset A_n$ with $\mu^*(G_n\setminus A_n)<\frac{\varepsilon}{2^n}$. Then $A\equiv\bigcup_{n=1}^{\infty}A_n$ is contained in the open set $G\equiv\bigcup_{n=1}^{\infty}G_n$, and since $G\setminus A$ is contained in $\bigcup_{n=1}^{\infty}(G_n\setminus A_n)$, monotonicity and subadditivity of μ^* yield

$$\mu^*(G\setminus A)\leq\mu^*\left(\bigcup_{n=1}^{\infty}(G_n\setminus A_n)\right)\leq\sum_{n=1}^{\infty}\mu^*(G_n\setminus A_n)<\sum_{n=1}^{\infty}\frac{\varepsilon}{2^n}=\varepsilon.$$

Step 3: $A\in\mathcal{L}$ if A is closed.

Suppose first that A is compact, and let $\varepsilon>0$. Then using Claim 2 there is $G=\dot{\bigcup}_{n=1}^{N^*}(a_n,b_n)$ containing A with

$$\mu^*(G)=\sum_{n=1}^{\infty}(b_n-a_n)\leq\mu^*(A)+\varepsilon<\infty.$$

Now $G\setminus A$ is open and so $G\setminus A=\dot{\bigcup}_{m=1}^{M^*}(c_m,d_m)$ by Proposition 14. We want to show that $\mu^*(G\setminus A)\leq\varepsilon$. Fix a finite $M\leq M^*$ and

$$0<\eta<\frac{1}{2}\min_{1\leq m\leq M}(d_m-c_m).$$

Then the compact set

$$K_\eta=\bigcup_{m=1}^{M}[c_m+\eta,d_m-\eta]$$

is disjoint from A, so by Claim 3 we have

$$\mu^*(A)+\mu^*(K_\eta)=\mu^*(A\cup K_\eta).$$

We conclude from subadditivity and $A \cup K_\eta \subset G$ that

$$\begin{aligned}\mu^*(A) + \sum_{m=1}^{M}(d_m - c_m - 2\eta) &= \mu^*(A) + \mu^*\left(\bigcup_{m=1}^{M}(c_m + \eta, d_m - \eta)\right) \\ &\leq \mu^*(A) + \mu^*(K_\eta) \\ &= \mu^*(A \cup K_\eta) \\ &\leq \mu^*(G) \leq \mu^*(A) + \varepsilon.\end{aligned}$$

Since $\mu^*(A) < \infty$ for A compact, we thus have

$$\sum_{m=1}^{M}(d_m - c_m) \leq \varepsilon + 2M\eta$$

for all $0 < \eta < \frac{1}{2}\min_{1 \leq m \leq M}(d_m - c_m)$. Hence $\sum_{m=1}^{M}(d_m - c_m) \leq \varepsilon$ and taking the supremum in $M \leq M^*$ we obtain from Claim 2 that

$$\mu^*(G \setminus A) = \sum_{m=1}^{M^*}(d_m - c_m) \leq \varepsilon.$$

Finally, if A is closed, it is a countable union of compact sets $A = \bigcup_{n=1}^{\infty}([-n, n] \cap A)$, and hence $A \in \mathcal{L}$ by Step 2.

Step 4: $A^c \in \mathcal{L}$ if $A \in \mathcal{L}$.

For each $n \geq 1$ there is by Claim 2 an open set $G_n \supset A$ such that $\mu^*(G_n \setminus A) < \frac{1}{n}$. Then $F_n \equiv G_n^c$ is closed and hence $F_n \in \mathcal{L}$ by Step 3. Thus

$$S \equiv \bigcup_{n=1}^{\infty} F_n \in \mathcal{L}, \quad S \subset A^c,$$

and $A^c \setminus S \subset G_n \setminus A$ for all n implies that

$$\mu^*(A^c \setminus S) \leq \mu^*(G_n \setminus A) < \frac{1}{n}, \qquad n \geq 1.$$

Thus $\mu^*(A^c \setminus S) = 0$ and by Step 1 we have $A^c \setminus S \in \mathcal{L}$. Finally, Step 2 shows that

$$A^c = S \cup (A^c \setminus S) \in \mathcal{L}.$$

Thus far we have shown that $\mathcal{L}$ is a σ-algebra, and we now turn to proving that $\mathcal{L}$ and μ satisfy the four properties in Theorem 51. Property (1) is an easy exercise. Property (2) is the main event. Let $\{E_n\}_{n=1}^{\infty}$ be a pairwise disjoint sequence of sets in $\mathcal{L}$, and let $E = \dot{\bigcup}_{n=1}^{\infty} E_n$.

We will consider first the case where each of the sets E_n is bounded. Let $\varepsilon > 0$ be given. Then $E_n^c \in \mathcal{L}$ and so there are open sets $G_n \supset E_n^c$ such that

$$\mu^*(G_n \setminus E_n^c) < \frac{\varepsilon}{2^n}, \qquad n \geq 1.$$

Equivalently, with $F_n = G_n^c$, we have F_n closed, contained in E_n, and

$$\mu^*(E_n \setminus F_n) < \frac{\varepsilon}{2^n}, \qquad n \geq 1.$$

Thus the sets $\{F_n\}_{n=1}^{\infty}$ are compact and pairwise disjoint. Claim 3 and induction shows that

$$\sum_{n=1}^{N} \mu^* (F_n) = \mu^* \left(\bigcup_{n=1}^{N} F_n \right) \leq \mu^* (E) , \qquad N \geq 1,$$

and taking the supremum over N yields

$$\sum_{n=1}^{\infty} \mu^* (F_n) \leq \mu^* (E) .$$

Thus we have

$$\begin{aligned} \sum_{n=1}^{\infty} \mu^* (E_n) &\leq \sum_{n=1}^{\infty} \{\mu^* (E_n \setminus F_n) + \mu^* (F_n)\} \\ &\leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} + \sum_{n=1}^{\infty} \mu^* (F_n) \leq \varepsilon + \mu^* (E) . \end{aligned}$$

Since $\varepsilon > 0$ we conclude that $\sum_{n=1}^{\infty} \mu^* (E_n) \leq \mu^* (E)$, and subadditivity of μ^* then proves equality.

In general, define $E_{n,k} = E_n \cap \{(-k-1, -k] \cup [k, k+1)\}$ for $k, n \geq 1$ so that

$$E = \dot{\bigcup}_{n=1}^{\infty} E_n = \dot{\bigcup}_{n,k=1}^{\infty} E_{n,k}.$$

Then from what we just proved we have

$$\mu^* (E) = \sum_{n,k=1}^{\infty} \mu^* (E_{n,k}) = \sum_{n=1}^{\infty} \left(\sum_{k=1}^{\infty} \mu^* (E_{n,k}) \right) = \sum_{n=1}^{\infty} \mu^* (E_n) .$$

Finally, property (3) follows from the observation that $E \subset \dot{\bigcup}_{n=1}^{\infty} (a_n, b_n)$ if and only if $E + x \subset \dot{\bigcup}_{n=1}^{\infty} (a_n + x, b_n + x)$. It is then obvious that $\mu^* (E + x) = \mu^* (E)$ and that $E + x \in \mathcal{L}$ if $E \in \mathcal{L}$. Property (4) is immediate from Step 1 above. This completes the proof of Theorem 51.

2. Measurable functions and integration

Let $[-\infty, \infty] = \mathbb{R} \cup \{-\infty, \infty\}$ be the extended real numbers with order and (some) algebra operations defined by

$$\begin{aligned} -\infty &< x < \infty, \qquad x \in \mathbb{R}, \\ x + \infty &= \infty, \qquad x \in \mathbb{R}, \\ x - \infty &= -\infty, \qquad x \in \mathbb{R}, \\ x \cdot \infty &= \infty, \qquad x > 0, \\ x \cdot \infty &= -\infty, \qquad x < 0, \\ 0 \cdot \infty &= 0. \end{aligned}$$

The final assertion $0 \cdot \infty = 0$ is dictated by $\sum_{n=1}^{\infty} a_n = 0$ if all the $a_n = 0$. It turns out that these definitions give rise to a consistent theory of measure and integration of functions with values in the extended real number system.

Let $f : \mathbb{R} \to [-\infty, \infty]$. We say that f is (Lebesgue) measurable if

$$f^{-1} ([-\infty, x)) \in \mathcal{L}, \qquad x \in \mathbb{R}.$$

The simplest examples of measurable functions are the characteristic functions χ_E of measurable sets E. Indeed,

$$\left(\chi_E\right)^{-1}\left([-\infty,x)\right)=\begin{cases}\phi & \text{if} \quad x\leq 0\\ E^c & \text{if} \quad 0<x\leq 1\\ \mathbb{R} & \text{if} \quad x>1\end{cases}.$$

It is then easy to see that finite linear combinations $s=\sum_{n=1}^{N}a_n\chi_{E_n}$ of such characteristic functions χ_{E_n}, called simple functions, are also measurable. Here $a_n\in\mathbb{R}$ and E_n is a measurable subset of $\mathbb{R}$. It turns out that if we define the integral of a simple function $s=\sum_{n=1}^{N}a_n\chi_{E_n}$ by

$$\int_{\mathbb{R}}s=\sum_{n=1}^{N}a_n\mu\left(E_n\right),$$

the value is independent of the representation of s as a simple function. Armed with this fact we can then extend the definition of integral $\int_{\mathbb{R}}f$ to functions f that are nonnegative on $\mathbb{R}$, and then to functions f such that $\int_{\mathbb{R}}|f|<\infty$.

At each stage one establishes the relevant properties of the integral along with the most useful theorems. For the most part these extensions are rather routine, the cleverness inherent in the theory being in the overarching organization of the concepts rather than in the details of the demonstrations. As a result, we will merely state the main results in logical order and sketch proofs when not simply routine. We will however give fairly detailed proofs of the three famous convergence theorems, the Monotone Convergence Theorem, Fatou's Lemma, and the Dominated Convergence Theorem. The reader is referred to the excellent exposition in [**5**] for the complete story including many additional fascinating insights.

2.1. Properties of measurable functions. From now on we denote the Lebesgue measure of a measurable subset E of $\mathbb{R}$ by $|E|$ rather than by $\mu(E)$ as in the previous sections. We say that two measurable functions $f,g:\mathbb{R}\to[-\infty,\infty]$ are equal *almost everywhere* (often abbreviated *a.e.*) if

$$|\{x\in\mathbb{R}:f(x)\neq g(x)\}|=0.$$

We say that f is *finite-valued* if $f:\mathbb{R}\to\mathbb{R}$. We now collect a number of elementary properties of measurable functions.

LEMMA 26. *Suppose that $f,f_n,g:\mathbb{R}\to[-\infty,\infty]$ for $n\in\mathbb{N}$.*

(1) *If f is finite-valued, then f is measurable* if and only if *$f^{-1}(G)\in\mathcal{L}$ for all open sets $G\subset\mathbb{R}$* if and only if *$f^{-1}(F)\in\mathcal{L}$ for all closed sets $F\subset\mathbb{R}$.*
(2) *If f is finite-valued and continuous, then f is measurable.*
(3) *If f is finite-valued and measurable and $\Phi:\mathbb{R}\to\mathbb{R}$ is continuous, then $\Phi\circ f$ is measurable.*
(4) *If $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions, then the following functions are all measurable:*
$$\sup_n f_n(x),\quad \inf_n f_n(x),\ldots\limsup_{n\to\infty} f_n(x),\quad \liminf_{n\to\infty} f_n(x).$$
(5) *If $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions and $f(x)=\lim_{n\to\infty}f_n(x)$, then f is measurable.*
(6) *If f is measurable, so is f^n for $n\in\mathbb{N}$.*
(7) *If f and g are finite-valued and measurable, then so are $f+g$ and fg.*

(8) *If f is measurable and $f = g$ almost everywhere, then g is measurable.*

Comments: For property (1), first show that f is measurable if and only if $f^{-1}((a,b)) \in \mathcal{L}$ for all $-\infty < a < b < \infty$. For property (3) use $(\Phi \circ f)^{-1}(G) = f^{-1}(\Phi^{-1}(G))$ and note that $\Phi^{-1}(G)$ is open if G is open. For property (7), use

$$\begin{aligned} \{f + g > a\} &= \bigcup_{r \in \mathbb{Q}} [\{f > a - r\} \cap \{g > r\}], \qquad a \in \mathbb{R}, \\ fg &= \frac{1}{4}\left[(f+g)^2 - (f-g)^2\right]. \end{aligned}$$

Recall that a measurable simple function φ (i.e. the range of φ is finite) has the form

$$\varphi = \sum_{k=1}^{N} \alpha_k \chi_{E_k}, \qquad \alpha_k \in \mathbb{R}, E_k \in \mathcal{L}.$$

Next we collect two approximation properties of simple functions.

PROPOSITION 18. *Let $f : \mathbb{R} \to [-\infty, \infty]$ be measurable.*

(1) *If f is nonnegative there is an increasing sequence of nonnegative simple functions $\{\varphi_k\}_{k=1}^{\infty}$ that converges pointwise and monotonically to f:*

$$\varphi_k(x) \le \varphi_{k+1}(x) \text{ and } \lim_{k\to\infty} \varphi_k(x) = f(x), \qquad \text{for all } x \in \mathbb{R}.$$

(2) *There is a sequence of simple functions $\{\varphi_k\}_{k=1}^{\infty}$ satisfying*

$$|\varphi_k(x)| \le |\varphi_{k+1}(x)| \text{ and } \lim_{k\to\infty} \varphi_k(x) = f(x), \qquad \text{for all } x \in \mathbb{R}.$$

Comments: To prove (1) let $f_M = \min\{f, M\}$, and for $0 \le n < NM$ define

$$E_{n,N,M} = \left\{x \in \mathbb{R} : \frac{n}{N} < f_M(x) \le \frac{n+1}{N}\right\}.$$

Then $\varphi_k(x) = \sum_{n=1}^{2^k k} \frac{n}{2^k} \chi_{E_{n,2^k,k}}(x)$ works. Property (2) is routine given (1).

2.2. Properties of integration and convergence theorems. If φ is a measurable simple function (i.e. its range is a finite set), then φ has a unique canonical representation

$$\varphi = \sum_{k=1}^{N} \alpha_k \chi_{E_k},$$

where the real constants α_k are distinct and nonzero, and the measurable sets E_k are pairwise disjoint. We define the Lebesgue integral of φ by

$$\int \varphi(x)\, dx = \sum_{k=1}^{N} \alpha_k |E_k|.$$

If E is a measurable subset of $\mathbb{R}$ and φ is a measurable simple function, then so is $\chi_E \varphi$, and we define

$$\int_E \varphi(x)\, dx = \int (\chi_E \varphi)(x)\, dx.$$

LEMMA 27. *Suppose that φ and ψ are measurable simple functions and that $E, F \in \mathcal{L}$.*

(1) *If* $\varphi = \sum_{k=1}^{M} \beta_k \chi_{F_k}$ *(not necessarily the canonical representation), then*

$$\int \varphi(x)\,dx = \sum_{k=1}^{M} \beta_k |F_k|.$$

(2) $\int (a\varphi + b\psi) = a \int \varphi + b \int \psi$ *for* $a, b \in \mathbb{C}$,
(3) $\int_{E \cup F} \varphi = \int_E \varphi + \int_F \varphi$ *if* $E \cap F = \phi$,
(4) $\int \varphi \leq \int \psi$ *if* $\varphi \leq \psi$,
(5) $\left|\int \varphi\right| \leq \int |\varphi|$.

Properties (2) - (5) are usually referred to as *linearity*, *additivity*, *monotonicity* and the *triangle inequality* respectively. The proofs are routine.

Now we turn to defining the integral of a nonnegative measurable function $f : \mathbb{R} \to [0, \infty]$. For such f we define

$$\int f(x)\,dx = \sup \left\{ \int g(x)\,dx : 0 \leq \varphi \leq f \text{ and } \varphi \text{ is simple} \right\}.$$

It is essential here that f be permitted to take on the value ∞, and that the supremum may be ∞ as well. We say that f is (Lebesgue) *integrable* if $\int f(x)\,dx < \infty$. For E measurable define

$$\int_E f(x)\,dx = \int (\chi_E f)(x)\,dx.$$

Here is an analogue of Lemma 27 whose proof is again routine.

LEMMA 28. *Suppose that* $f, g : \mathbb{R} \to [0, \infty]$ *are nonnegative measurable functions and that* $E, F \in \mathcal{L}$.

(1) $\int (af + bg) = a \int f + b \int g$ *for* $a, b \in (0, \infty)$,
(2) $\int_{E \cup F} f = \int_E f + \int_F f$ *if* $E \cap F = \phi$,
(3) $\int f \leq \int g$ *if* $0 \leq f \leq g$,
(4) *If* $\int f < \infty$, *then* $f(x) < \infty$ *for a.e.* x,
(5) *If* $\int f = 0$, *then* $f(x) = 0$ *for a.e.* x.

Note that convergence of integrals does not always follow from pointwise convergence of the integrands. For example,

$$\lim_{n\to\infty} \int \chi_{[n,n+1]}(x)\,dx = 1 \neq 0 = \int \lim_{n\to\infty} \chi_{[n,n+1]}(x)\,dx,$$

and

$$\lim_{n\to\infty} \int n\chi_{\left(0,\frac{1}{n}\right)}(x)\,dx = 1 \neq 0 = \int \lim_{n\to\infty} n\chi_{\left[0,\frac{1}{n}\right]}(x)\,dx.$$

In each of these examples, the mass of the integrands "disappears" in the limit; at "infinity" in the first example and at the origin in the second example. Here are our first two classical convergence theorems giving conditions under which convergence *does* hold.

THEOREM 52. *(Monotone Convergence Theorem) Suppose that* $\{f_n\}_{n=1}^{\infty}$ *is an* increasing *sequence of nonnegative measurable functions, i.e.* $f_n(x) \leq f_{n+1}(x)$, *and let*

$$f(x) = \sup_n f_n(x) = \lim_{n\to\infty} f_n(x).$$

Then f is nonegative and measurable and

$$\int f(x)\,dx = \lim_{n\to\infty}\int f_n(x)\,dx.$$

Proof: Since $\int f_n \le \int f_{n+1}$ we have $\lim_{n\to\infty}\int f_n = L \in [0,\infty]$. Now f is measurable and $f_n \le f$ implies $\int f_n \le \int f$ so that

$$L \le \int f.$$

To prove the opposite inequality, momentarily fix a simple function φ such that $0 \le \varphi \le f$. Choose $c < 1$ and define

$$E_n = \{x \in \mathbb{R} : f_n(x) \ge c\varphi(x)\}, \qquad n \ge 1.$$

Then E_n is an increasing sequence of measurable sets with $\bigcup_{n=1}^{\infty} E_n = \mathbb{R}$. We have

$$\int f_n \ge \int_{E_n} f_n \ge c\int_{E_n}\varphi, \qquad n \ge 1.$$

Now let $\varphi = \sum_{k=1}^{N}\alpha_k\chi_{F_k}$ be the canonical representation of φ. Then

$$\int_{E_n}\varphi = \sum_{k=1}^{N}\alpha_k\,|E_n \cap F_k|,$$

and since $\lim_{n\to\infty}|E_n \cap F_k| = |F_k|$ by the fourth line in (1.1), we obtain that

$$\int_{E_n}\varphi = \sum_{k=1}^{N}\alpha_k\,|E_n \cap F_k| \to \sum_{k=1}^{N}\alpha_k\,|F_k| = \int\varphi$$

as $n \to \infty$. Altogether then we have

$$L = \lim_{n\to\infty}\int f_n \ge c\int\varphi$$

for all $c < 1$, which implies $L \ge \int\varphi$ for all simple φ with $0 \le \varphi \le f$, which implies $L \ge \int f$ as required.

Corollary 14. *Suppose that $a_k(x) \ge 0$ is measurable for $k \ge 1$. Then*

$$\int\sum_{k=1}^{\infty}a_k(x)\,dx = \sum_{k=1}^{\infty}\int a_k(x)\,dx.$$

To prove the corollary apply the Monotone Convergence Theorem to the sequence of partial sums $f_n(x) = \sum_{k=1}^{n}a_k(x)$.

Lemma 29. *(Fatou's Lemma) If $\{f_n\}_{n=1}^{\infty}$ is a sequence of* nonnegative *measurable functions, then*

$$\int\liminf_{n\to\infty} f_n(x)\,dx \le \liminf_{n\to\infty}\int f_n(x)\,dx.$$

Proof: Let $g_n(x) = \inf_{k\ge n} f_k(x)$ so that $g_n \le f_n$ and $\int g_n \le \int f_n$. Then $\{g_n\}_{n=1}^{\infty}$ is an increasing sequence of nonnegative measurable functions that converges pointwise to $\liminf_{n\to\infty} f_n(x)$. So the Monotone Convergence Theorem yields

$$\int\liminf_{n\to\infty} f_n(x)\,dx = \lim_{n\to\infty}\int g_n(x)\,dx \le \liminf_{n\to\infty}\int f_n(x)\,dx.$$

Finally, we can give an unambiguous meaning to the integral $\int f(x)\,dx$ in the case when f is *integrable*, by which we mean that f is measurable and $\int |f(x)|\,dx < \infty$. To do this we introduce the positive and negative parts of f:

$$f^{+}(x) = \max\{f(x), 0\} \text{ and } f_{-}(x) = \max\{-f(x), 0\}.$$

Then both f^{+} and f_{-} are nonnegative measurable functions with finite integral. We define

$$\int f(x)\,dx = \int f^{+}(x)\,dx - \int f_{-}(x)\,dx.$$

With this definition we have the usual elementary properties of linearity, additivity, monotonicity and the triangle inequality.

LEMMA 30. *Suppose that f, g are integrable and that $E, F \in \mathcal{L}$.*

1. $\int (af + bg) = a \int f + b \int g$ *for* $a, b \in \mathbb{R}$,
2. $\int_{E \cup F} f = \int_E f + \int_F f$ *if* $E \cap F = \phi$,
3. $\int f \le \int g$ *if* $f \le g$,
4. $\left|\int f\right| \le \int |f|$.

Our final convergence theorem is one of the most useful in analysis.

THEOREM 53. *(Dominated Convergence Theorem) Let g be a nonnegative integrable function. Suppose that $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions satisfying*

$$\lim_{n\to\infty} f_n(x) = f(x), \qquad a.e.\ x,$$

and

$$|f_n(x)| \le g(x), \qquad a.e.\ x.$$

Then

$$\lim_{n\to\infty} \int |f(x) - f_n(x)|\,dx = 0,$$

and hence

$$\int f(x)\,dx = \lim_{n\to\infty} \int f_n(x)\,dx.$$

Proof: Since $|f| \le g$ and f is measurable, f is integrable. Since $|f - f_n| \le 2g$, Fatou's Lemma can be applied to the sequence of functions $2g - |f - f_n|$ to obtain

$$\begin{aligned} \int 2g &\le \liminf_{n\to\infty} \int (2g - |f - f_n|) \\ &= \int 2g + \liminf_{n\to\infty} \left(-\int |f - f_n|\right) \\ &= \int 2g - \limsup_{n\to\infty} \int |f - f_n|. \end{aligned}$$

Since $\int 2g < \infty$, we can subtract it from both sides to obtain

$$\limsup_{n\to\infty} \int |f - f_n| \le 0,$$

which implies $\lim_{n\to\infty} \int |f - f_n| = 0$. Then $\int f = \lim_{n\to\infty} \int f_n$ follows from the triangle inequality $\left|\int (f - f_n)\right| \le \int |f - f_n|$.

Finally, if $f(x) = u(x) + iv(x)$ is complex-valued where $u(x)$ and $v(x)$ are real-valued measurable functions such that

$$\int |f(x)|\, dx = \int \sqrt{u(x)^2 + v(x)^2} dx < \infty,$$

then we define

$$\int f(x)\, dx = \int u(x)\, dx + i \int v(x)\, dx.$$

The usual properties of linearity, additivity, monotonicity and the triangle inequality all hold for this definition as well.

2.3. Three famous measure problems. The following three problems are listed in order of increasing difficulty.

PROBLEM 3. *Suppose that* $E_1, ..., E_n$ *are* n *Lebesgue measurable subsets of* $[0,1]$ *such that each point* x *in* $[0,1]$ *lies in some* k *of these subsets. Prove that there is at least one set* E_j *with* $|E_j| \geq \frac{k}{n}$.

PROBLEM 4. *Suppose that* E *is a Lebesgue measurable set of positive measure. Prove that*

$$E - E = \{x - y : x, y \in E\}$$

contains a nontrivial open interval.

PROBLEM 5. *Construct a Lebesgue measurable subset of the real line such that*

$$0 < \frac{|E \cap I|}{|I|} < 1$$

for all nontrivial open intervals I.

To solve Problem 3, note that the hypothesis implies $k \leq \sum_{j=1}^{n} \chi_{E_j}(x)$ for $x \in [0,1]$. Now integrate to obtain

$$k = \int_0^1 k dx \leq \int_0^1 \left(\sum_{j=1}^{n} \chi_{E_j}(x) \right) dx = \sum_{j=1}^{n} \int_0^1 \chi_{E_j}(x)\, dx = \sum_{j=1}^{n} |E_j|,$$

which implies that $|E_j| \geq \frac{k}{n}$ for some j. The solution is much less elegant without recourse to integration.

To solve Problem 4, choose K compact contained in E such that $|K| > 0$. Then choose G open containing K such that $|G \setminus K| < |K|$. Let $\delta = dist(K, G^c) > 0$. It follows that $(-\delta, \delta) \subset K - K \subset E - E$. Indeed, if $x \in (-\delta, \delta)$ then $K - x \subset G$ and $K \cap (K - x) \neq \phi$ since otherwise we have a contradiction:

$$2|K| = |K| + |K - x| \leq |G| \leq |G \setminus K| + |K| < 2|K|.$$

Thus there are k_1 and k_2 in K such that $k_1 = k_2 - x$ and so

$$x = k_2 - k_1 \in K - K.$$

Problem 5 is most easily solved using generalized Cantor sets E_α. Let $0 < \alpha \leq 1$ and set $I_1^0 = [0,1]$. Remove the open interval of length $\frac{1}{3}\alpha$ centered in I_1^0 and denote the two remaining closed intervals by I_1^1 and I_2^1. Then remove the open interval of length $\frac{1}{3^2}\alpha$ centered in I_1^1 and denote the two remaining closed intervals by I_1^2 and I_2^2. Do the same for I_2^1 and denote the two remaining closed intervals by I_3^2 and I_4^2.

Continuing in this way, we obtain at the k^{th} generation, a collection $\left\{I_j^k\right\}_{j=1}^{2^k}$ of 2^k pairwise disjoint closed intervals of equal length. Let

$$E_\alpha = \bigcap_{k=1}^{\infty} \left(\bigcup_{j=1}^{2^k} I_j^k \right).$$

Then by summing the lengths of the removed open intervals, we obtain

$$|[0,1] \setminus E_\alpha| = \frac{1}{3}\alpha + \frac{2}{3^2}\alpha + \frac{2^2}{3^3}\alpha + ... = \alpha,$$

and it follows that E_α is compact and has Lebesgue measure $1-\alpha$. It is not hard to show that E_α is also nowhere dense. The case $\alpha = 1$ is particularly striking: E_1 is a compact, perfect and uncountable subset of $[0,1]$ having Lebesgue measure 0. This is the classical Cantor set introduced as a fractal in Subsection 3.1 of Chapter 3.

In order to construct the set E in Problem 3, it suffices by taking unions of translates by integers, to construct a subset E of $[0,1]$ satisfying

$$0 < \frac{|E \cap I|}{|I|} < 1, \qquad \text{for all intervals } I \subset [0,1] \text{ of positive length.} \tag{2.1}$$

Fix $0 < \alpha_1 < 1$ and start by taking $E^1 = E_{\alpha_1}$. It is not hard to see that $\frac{|E^1 \cap I|}{|I|} < 1$ for all I, but the left hand inequality in (2.1) fails for $E = E^1$ whenever I is a subset of one of the component intervals in the open complement $[0,1] \setminus E^1$. To remedy this fix $0 < \alpha_2 < 1$ and for each component interval J of $[0,1] \setminus E^1$, translate and dilate E_{α_2} to fit snugly in the closure $\overline{J}$ of the component, and let E^2 be the union of E^1 and all these translates and dilates of E_{α_2}. Then again, $\frac{|E^2 \cap I|}{|I|} < 1$ for all I but the left hand inequality in (2.1) fails for $E = E^2$ whenever I is a subset of one of the component intervals in the open complement $[0,1] \setminus E^2$. Continue this process indefinitely with a sequence of numbers $\{\alpha_n\}_{n=1}^{\infty} \subset (0,1)$. We claim that $E = \bigcup_{n=1}^{\infty} E^n$ satisfies (2.1) if and only if

$$\sum_{n=1}^{\infty} (1 - \alpha_n) < \infty. \tag{2.2}$$

To see this, first note that no matter what sequence of numbers α_n *less than one* is used, we obtain that $0 < \frac{|E \cap I|}{|I|}$ for all intervals I of positive length. Indeed, each set E^n is easily seen to be compact and nowhere dense, and each component interval in the complement $[0,1] \setminus E^n$ has length at most

$$\frac{\alpha_1}{3} \frac{\alpha_2}{3} ... \frac{\alpha_n}{3} \leq 3^{-n}.$$

Thus given an interval I of positive length, there is n large enough such that I will contain one of the component intervals J of $[0,1] \setminus E^n$, and hence will contain the translated and dilated copy $\mathcal{C}\left(E_{\alpha_{n+1}}\right)$ of $E_{\alpha_{n+1}}$ that is fitted into J by construction. Since the dilation factor is the length $|J|$ of J, we have

$$|E \cap I| \geq \left|\mathcal{C}\left(E_{\alpha_{n+1}}\right)\right| = |J| \left|E_{\alpha_{n+1}}\right| = |J| (1 - \alpha_{n+1}) > 0,$$

since $\alpha_{n+1} < 1$.

It remains to show that $|E \cap I| < |I|$ for all intervals I of positive length in $[0,1]$, and it is here that we must use (2.2). Indeed, fix I and let J be a component interval of $[0,1] \setminus E^n$ (with n large) that is contained in I. Let $\mathcal{C}\left(E_{\alpha_{n+1}}\right)$ be the translated and dilated copy of $E_{\alpha_{n+1}}$ that is fitted into J by construction. We compute that

$$\begin{aligned} |E \cap J| &= \left|\mathcal{C}\left(E_{\alpha_{n+1}}\right)\right| + (1-\alpha_{n+2}) \left|J \setminus \mathcal{C}\left(E_{\alpha_{n+1}}\right)\right| + ... \\ &= (1-\alpha_{n+1}) |J| + (1-\alpha_{n+2}) (1-(1-\alpha_{n+1})) |J| \\ &\quad + (1-\alpha_{n+3}) (1-(1-\alpha_{n+1}) - (1-\alpha_{n+2}) (1-(1-\alpha_{n+1}))) |J| + ... \\ &= \sum_{k=1}^{\infty} \beta_k^n |J|, \end{aligned}$$

where by induction,

$$\beta_k^n = (1-\alpha_{n+k}) \alpha_{n+k-1} ... \alpha_{n+1}, \qquad k \geq 1.$$

Then we have

$$|E \cap J| = \left(\sum_{k=1}^{\infty} \beta_k^n\right) |J| < |J|,$$

and hence also $\frac{|E \cap I|}{|I|} < 1$, if we choose $\{\alpha_n\}_{n=1}^{\infty}$ so that $\sum_{k=1}^{\infty} \beta_k^n < 1$ for all n.

Now we have

$$\sum_{k=1}^{\infty} \beta_k^n = \sum_{k=1}^{\infty} (1-\alpha_{n+k}) \alpha_{n+k-1} ... \alpha_{n+1} = 1 - \prod_{k=1}^{\infty} \alpha_{n+k},$$

and by the first line in (2.3) below, this is strictly less than 1 *if and only if* $\sum_{k=1}^{\infty} (1-\alpha_{n+k}) < \infty$ for all n. Thus the set E constructed above satisfies (2.1) if and only if (2.2) holds.

2.3.1. *Infinite products.* If $0 \leq u_n < 1$ and $0 \leq v_n < \infty$ then

$$\begin{aligned} \prod_{n=1}^{\infty} (1-u_n) &> 0 \text{ if and only if } \sum_{n=1}^{\infty} u_n < \infty, \\ \prod_{n=1}^{\infty} (1+v_n) &< \infty \text{ if and only if } \sum_{n=1}^{\infty} v_n < \infty. \end{aligned} \tag{2.3}$$

To see (2.3) we may assume $0 \leq u_n, v_n \leq \frac{1}{2}$, so that $e^{-u_n} \geq 1-u_n \geq e^{-2u_n}$ and $e^{\frac{1}{2}v_n} \leq 1+v_n \leq e^{v_n}$. For example, when $0 \leq x \leq \frac{1}{2}$, the alternating series estimate yields

$$e^{-2x} \leq 1-2x+\frac{(2x)^2}{2!} \leq 1-x,$$

while the geometric series estimate yields

$$e^{\frac{1}{2}x} \leq 1+\left(\frac{1}{2}x\right)\left\{1+x+x^2+...\right\} \leq 1+x.$$

Thus we have

$$\begin{aligned} \exp\left(-\sum_{n=1}^{\infty} u_n\right) &\geq \prod_{n=1}^{\infty} (1-u_n) \geq \exp\left(-2\sum_{n=1}^{\infty} u_n\right), \\ \exp\left(\frac{1}{2}\sum_{n=1}^{\infty} v_n\right) &\leq \prod_{n=1}^{\infty} (1+v_n) \leq \exp\left(\sum_{n=1}^{\infty} v_n\right). \end{aligned} \tag{2.4}$$

Bibliography

[1] R. G. Bartle and D. R. Sherbert, *Introduction to Real Analysis*, John Wiley and Sons, Inc. 3rd edition, 2000.

[2] C. B. Boyer, *A history of mathematics*, John Wiley & Sons, Inc., 1968.

[3] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, 3rd edition, 1976.

[4] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, 3rd edition, 1987.

[5] E. M. Stein and R. Shakarchi, *Complex Analysis, Princeton Lectures in Analysis* **II**, Princeton University Press, Princeton and Oxford, 2003.

[6] E. M. Stein and R. Shakarchi, *Real Analysis, Princeton Lectures in Analysis* **III**, Princeton University Press, Princeton and Oxford, 2003.

[7] S. Wagon, *The Banach-Tarski Paradox*, Cambridge University Press, 1985.