Daniel W. Stroock

A Concise Introduction to Analysis

Springer

A Concise Introduction to Analysis

Daniel W. Stroock

A Concise Introduction to Analysis

Springer

Daniel W. Stroock
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA
USA

ISBN 978-3-319-24467-9 ISBN 978-3-319-24469-3 (eBook)
DOI 10.1007/978-3-319-24469-3

Library of Congress Control Number: 2015950884

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

This book is dedicated to Elliot Gorokhovsky

Preface

This book started out as a set of notes that I wrote for a freshman high school student named Elliot Gorokhovsky. I had volunteered to teach an informal seminar at Fairview High School in Boulder, Colorado, but when it became known that participation in the seminar would not be for credit and would not appear on any transcript, the only student who turned up was Elliot.

Initially I thought that I could introduce Elliot to probability theory, but it soon became clear that my own incompetence in combinatorics combined with his ignorance of analysis meant that we would quickly run out of material. Thus I proposed that I teach him some of the analysis that is required to understand non-combinatorial probability theory. With this in mind, I got him a copy of Courant's venerable *Differential and Integral Calculus*. However, I felt that I ought to supplement Courant's book with some notes that presented the same material from a more modern perspective. Rudin's superb *Principles of Mathematical Analysis* provides an excellent introduction to the material that I wanted Elliot to learn, but it is too abstract for even a gifted high school freshman. I wanted Elliot to see a rigorous treatment of the fundamental ideas on which analysis is built, but I did not think that he needed to see them couched in great generality. Thus I ended up writing a book that would be a hybrid: part Courant and part Rudin and a little of neither. Perhaps the most closely related book is J. Marsden and M. Hoffman's *Elements of Classical Analysis*, although they cover a good deal more differential geometry and do not discuss complex analysis.

The book starts with differential calculus, first for functions of one variable (both real and complex). To provide the reader with interesting, concrete examples, thorough treatments are given of the trigonometric, exponential, and logarithmic functions. The next topic is the integral calculus for functions of a real variable, and, because it is more intuitive than Lebesgue's, I chose to develop Riemann's approach. The rest of the book is devoted to the differential and integral calculus for functions of more than one variable. Here too I use Riemann's integration theory, although I spend more time than is usually allotted to questions that smack of Lebesgue's theory. Prominent space is given to polar coordinates and the divergence theorem, and these are applied in the final chapter to a derivation of Cauchy's integral formula. Several applications of Cauchy's formula are given, although in no sense is my treatment of analytic function theory comprehensive. Instead, I have tried to whet the appetite of my readers so that they will want to find out more. As an inducement, in the final section I present D.J. Newman's "simple" proof of the prime number theorem. If that fails to serve as an enticement, nothing will.

As its title implies, the book is concise, perhaps to a degree that some may feel makes it terse. In part my decision to make it brief was a reaction to the flaccid, 800 page calculus texts that are commonly adopted. A second motivation came from my own teaching experience. At least in America, for many, if not most, of the students who study mathematics, the language in which it is taught is not (to paraphrase H. Weyl) the language that was sung to them in their cradles. Such students do not profit from, and often do not even attempt to read, the lengthy explanations with which those 800 pages are filled. For them a book that errs on the side of brevity is preferable to one that errs on the side of verbosity. Be that as it may, I hope that this book will be valuable to people who have some facility in mathematics and who either never had a calculus course or were not satisfied by the one that they had. It does not replace existing texts, but it may be a welcome supplement and, to some extent, an antidote to them. If a few others approach it in the same spirit and with the same joy as Elliot did, I will consider it a success.

Finally, I would be remiss to not thank Peter Landweber for his meticulous reading of the multiple incarnations of my manuscript. Peter is a respected algebraic topologist who, now that he has retired, is broadening his mathematical expertise. Both I and my readers are the beneficiaries of his efforts.

Daniel W. Stroock

Contents

Notations

$\equiv$	When this is used between two expressions, it means that the expression on the left is defined by the one of the right
$\mathrm{card}(S)$	The cardinality of (i.e., number of elements in) the set S
$\mathbb{R}$	The real line
$\mathbb{C}$	The complex plane of numbers $z = x + iy$, where $x, y \in \mathbb{R}$ and $i = \sqrt{-1}$
$\mathbb{R}^N$	N-dimensional Euclidean space
$B(\mathbf{c}, r)$	The open ball $\{\mathbf{x} \in \mathbb{R}^N : \lvert\mathbf{x} - \mathbf{c}\rvert < r\}$ in $\mathbb{R}^N$
$D(z, r)$	The open disk $\{\zeta \in \mathbb{C} : \lvert\zeta - z\rvert < r\}$ in $\mathbb{C}$
$\mathbb{S}^{N-1}(\mathbf{c}, r)$	The sphere $\{\mathbf{x} : \lvert\mathbf{x} - \mathbf{c}\rvert = r\}$ in $\mathbb{R}^N$
$(\mathbf{x}, \mathbf{y})_{\mathbb{R}^N}$	The inner product $\sum_{j=1}^N x_j y_j$ of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$
$a \vee b$	The larger of the numbers a and b
$a \wedge b$	The smaller of the numbers a and b
a^+ & a^-	The positive part $a \vee 0$ and negative part $-(a \wedge 0)$ of $a \in \mathbb{R}$
$\mathbb{Z}$	The set of all integers
$\mathbb{Z}^+$	The set of all positive integers
$\mathbb{N}$	The set of all non-negative integers (i.e., $Z^+ \cup \{0\}$)
Ω_N	The volume of the unit ball in $\mathbb{R}^N$
$\sum_{m=1}^n a_m$	The sum of the numbers $a_1, \ldots, a_n$
$\prod_{m=1}^n a_m$	The product of the numbers $a_1, \ldots, a_n$
$\overline{\lim}_{n\to\infty} a_n$	The limit superior of $\{a_n : n \geq 1\}$
$\underline{\lim}_{n\to\infty} a_n$	The limit inferior of $\{a_n : n \geq 1\}$
$S\complement$	The complement of the set S
$\mathrm{int}(S)$	The interior of the set S
$\bar{S}$	The closure of the set S
∂S	The boundary of the set S
$\mathfrak{R}(z)$ & $\mathfrak{I}(z)$	The real part x and the imaginary part y of the complex number $z = x + iy \in \mathbb{C}$
$f^{-1}(S)$	The inverse image $\{x : f(x) \in S\}$ of the set S under the function f

$f\restriction S$	The restriction of the function f to the set S
f'	The first derivative of the function f
$\dot{\mathbf{p}}$	The time derivative of a path $\mathbf{p}$
f'_φ	The directional derivative in the direction $e^{i\varphi}$ of the function f on $\mathbb{C}$
$\partial_\xi f$	The directional derivative in the direction ξ of the function f on $\mathbb{R}^N$
$f^{(m)} = \partial^m f$	The mth derivative of the function f
∇f	The gradient $(\partial_{\mathbf{e}_1} f, \ldots, \partial_{\mathbf{e}_N} f)$ of the function f
div$\mathbf{F}$	The divergence $\sum_{j=1}^N \partial_{\mathbf{e}_j} F_j$ of the $\mathbb{R}^N$-valued function $\mathbf{F} = (F_1, \ldots, F_N)$
$g \circ f$	The composition of the function g with the function f
$\|f\|_S$	The uniform norm $\sup\{\lvert f(x)\rvert : x \in S\}$ of a function f on the set S
$\mathfrak{R}(f; \mathcal{C}, \Xi)$	The Riemann sum $\sum_{I\in\mathcal{C}} f(\Xi(I))\lvert I\rvert$
$\mathfrak{e}_m(x)$	The imaginary exponential $e^{i2\pi mx}$
$\hat{f}_m$	The Fourier coefficient $\int_0^1 f(x)\mathfrak{e}_{-m}(x)\,dx$ of f
$\mathbf{1}_S(x)$	The indicator function of S, which is 1 if $x \in S$ and 0 otherwise

Chapter 1
Analysis on the Real Line

1.1 Convergence of Sequences in the Real Line

Calculus has developed systematic procedures for handling problems in which relationships between objects evolve and become exact only after one passes to a limit. As a consequence, one often ends up proving equalities between objects by what looks at first like the absurd procedure of showing that the difference between them is arbitrarily small. For example, consider the sequence of numbers $\frac{1}{n}$ for integers $n \geq 1$. Clearly, $\frac{1}{n}$ is not 0 for any integer. On the other hand, the larger n gets, the closer $\frac{1}{n}$ is to 0. To describe this in a mathematically precise way, one says $\frac{1}{n}$ converges to 0, by which one means that for any number $\epsilon > 0$ there is an integer n_ϵ such that $\left|\frac{1}{n}\right| < \epsilon$ for all $n \geq n_\epsilon$. More generally, given a sequence $\{x_n : n \geq 1\} \subseteq \mathbb{R}$,[1] one says that $\{x_n : n \geq 1\}$ *converges* in $\mathbb{R}$ if there is an $x \in \mathbb{R}$ with the property that for all $\epsilon > 0$ there exists an $n_\epsilon \in \mathbb{Z}^+$ such that $|x_n - x| < \epsilon$ whenever $n \geq n_\epsilon$, in which case one says that $\{x_n : n \geq 1\}$ *converges to* x and writes $x = \lim_{n\to\infty} x_n$ or $x_n \to x$. If for all $R > 0$ there exists an n_R such that $x_n \geq R$ for all $n \geq n_R$, then one says that $\{x_n : n \geq 1\}$ converges to ∞ and writes $\lim_{n\to\infty} x_n = \infty$ or $x_n \to \infty$. Similarly, if $\{-x_n : n \geq 1\}$ converges to ∞, then one says that $\{x_n : n \geq 1\}$ converges to $-\infty$ and writes $\lim_{n\to\infty} x_n = -\infty$ or $x_n \to -\infty$. Until one gets accustomed to this sort of thinking, it can be disturbing that the line of reasoning used to prove convergence often leads to a conclusion like $|x_n - x| \leq 5\epsilon$ for sufficiently large n's. However, as long as this conclusion holds for *all* $\epsilon > 0$, it should be clear that the presence of 5 or any other finite number makes no difference.

Given sequences $\{x_n : n \geq 1\}$ and $\{y_n : n \geq 1\}$ that converge, respectively, to $x \in \mathbb{R}$ and $y \in \mathbb{R}$, one can easily check that $\lim_{n\to\infty}(\alpha x_n + \beta y_n) = \alpha x + \beta y$ for all $\alpha, \beta \in \mathbb{R}$ and that $\lim_{n\to\infty} x_n y_n = xy$. To prove the first of these, simply observe that

[1] We will use $\mathbb{Z}$ to denote the set of all integers, $\mathbb{N}$ to denote the set of non-negative integers, and $\mathbb{Z}^+$ to denote the set of positive integers. The symbol $\mathbb{R}$ denotes the set of all real numbers. Also, we will use set theoretic notation to denote sequences even though a sequence should be thought of as a function on its index set.

D.W. Stroock, *A Concise Introduction to Analysis*,
DOI 10.1007/978-3-319-24469-3_1

$$\left|(\alpha x_n + \beta y_n) - (\alpha x + \beta y)\right| \le |\alpha||x_n - x| + |\beta||y_n - y| \longrightarrow 0 \quad \text{as } n \to \infty.$$

In the preceding, we used the *triangle inequality*, which is the easily verified statement that $|a + b| \le |a| + |b|$ for any pair of real numbers a and b. To prove the second, begin by noting that, because $|y_n| = |y + (y_n - y)| \le |y| + |y_n - y| \le |y| + 1$ for large enough n's, there is an $M < \infty$ such that $|y_n| \le M$ for all $n \ge 1$. Hence,

$$\begin{aligned}|x_n y_n - xy| &= |(x_n - x)y_n + x(y_n - y)| \le |x_n - x||y_n| + |x||y_n - y| \\ &\le M|x_n - x| + |x||y_n - y| \longrightarrow 0 \quad \text{as } n \to \infty.\end{aligned}$$

In both these we have used a frequently employed trick before applying the triangle inequality. Namely, because we wanted to relate the size of y_n to that of y and the difference between y_n and y, we added and subtracted y before applying the triangle inequality. Similarly, because we wanted to estimate the size of $x_n y_n - xy$ in terms of sizes of $x_n - x$ and $y_n - y$, it was convenient to add and subtract xy_n. Finally,

$$x_n \longrightarrow x \ne 0 \implies \frac{1}{x_n} \longrightarrow \frac{1}{x}.$$

Indeed, choose m so that $|x_n - x| \le \frac{|x|}{2}$ for $n \ge m$. Then

$$|x_n| + \frac{|x|}{2} \ge |x_n| + |x_n - x| \ge |x|,$$

and so $|x_n| \ge \frac{|x|}{2}$ for $n \ge m$. Hence

$$\left|\frac{1}{x_n} - \frac{1}{x}\right| = \frac{|x_n - x|}{|x_n||x|} \le \frac{2|x_n - x|}{|x|^2} \quad \text{for } n \ge m,$$

and so $\frac{1}{x_n} \longrightarrow \frac{1}{x}$.

Notice that if $x_n \to x \in \mathbb{R}$ and n_ϵ is chosen for $\epsilon > 0$ as above, then

$$|x_n - x_m| \le |x_n - x| + |x - x_m| \le 2\epsilon \quad \text{for } m, n \ge n_\epsilon.$$

That is, if $\{x_n : n \ge 1\}$ converges to some $x \in \mathbb{R}$, then the x_n's must be getting arbitrarily close to one another as $n \to \infty$. Now suppose that $\{x_n : n \ge 1\}$ is a sequence whose members are getting arbitrarily close to one another in the preceding sense. Then there are two possibilities. The first is that there is an empty hole in $\mathbb{R}$ and the x_n's are all crowding around it, in which case there would be nothing for them to converge to. The second possibility is that there are no holes in $\mathbb{R}$ and therefore that x_n's cannot be getting arbitrarily close to one another unless there is an $x \in \mathbb{R}$ to which they are converging. By giving a precise description of the real numbers, one can show that no holes exist, but, because it would be distracting to give such a description here, it has been deferred to the Appendix. For now we will simply accept the consequence of there being no holes in $\mathbb{R}$. That is, we will accept the statement, known as *Cauchy's convergence criterion*, that there is an $x \in \mathbb{R}$ to which $\{x_n : n \geq 1\}$ converges if, for each $\epsilon > 0$, there is an $n_e \geq 1$ such that $|x_{n'} - x_n| \leq \epsilon$ whenever $n, n' \geq n_\epsilon$. A space with a notion of convergence for which Cauchy's criterion guarantees convergence is said to be *complete*. Notice that, because $|x_{n'} - x_n| \leq |x_{n'} - x_m| + |x_n - x_m|$, $\{x_n : n \geq 1\}$ satisfies Cauchy's criterion if and only if, for each $\epsilon > 0$, there is an m such that $|x_n - x_m| < \epsilon$ for all $n > m$.

As a consequence of Cauchy's criterion, it is easy to show that a non-decreasing sequence $\{x_n : n \geq 1\}$ is convergent in $\mathbb{R}$ if it is bounded above (i.e., there is a $C < \infty$ such that $x_n \leq x_{n+1} \leq C$ for all $n \geq 1$). Indeed, suppose that it didn't and therefore that there exists an $\epsilon > 0$ with the property that for every m there is an $n > m$ such that $x_n - x_m \geq \epsilon$. Then we could choose $1 = n_1 < n_2 < \cdots < n_k < \cdots$ so that $x_{n_{k+1}} - x_{n_k} \geq \epsilon$ for all k, which would mean that $C \geq x_{n_k} \geq x_{n_1} + k\epsilon$ for all k, which is impossible. Similarly, if $\{x_n : n \geq 1\}$ is non-increasing and bounded below (i.e., there is a $C < \infty$ such that $-C \leq x_{n+1} \leq x_n$ for all n), then $\{x_n : n \geq 1\}$ converges in $\mathbb{R}$. In fact, this follows from the preceding by considering $\{-x_n : n \geq 1\}$. Finally, if $\{x_n : n \geq 1\}$ is non-decreasing and not bounded above, then $x_n \to \infty$, and if it is non-increasing and not bounded below, then $x_n \to -\infty$.

If $\{x_n : n \geq 1\}$ is a sequence and $1 \leq n_1 < \cdots < n_k < \cdots$, then we say that $\{x_{n_k} : k \geq 1\}$ is a *subsequence* of $\{x_n : n \geq 1\}$. It should be clear that if $\{x_n : n \geq 1\}$ converges in $\mathbb{R}$ or to $\pm\infty$, then so does every subsequence of $\{x_n : n \geq 1\}$. On the other hand, even if $\{x_n : n \geq 1\}$ doesn't converge, nonetheless it may admit subsequences that do. For instance, if $x_n = (-1)^n$, then $x_{2n-1} \to -1$ and $x_{2n} \to 1$. More generally, as will be shown in Theorem 1.3.3 below, if $\{x_n : n \geq 1\}$ is bounded (i.e., there is a $C < \infty$ such that $|x_n| \leq C$ for all n), then $\{x_n : n \geq 1\}$ admits a convergent subsequence. Finally, x is said to be a *limit point* of $\{x_n : n \geq 1\}$ if there is a subsequence that converges to x.

1.2 Convergence of Series

One often needs to sum an infinite number of real numbers. That is, given a sequence $\{a_m : m \geq 0\} \subseteq \mathbb{R}$, one would like to assign a meaning to

$$\sum_{m=0}^{\infty} a_m = a_0 + \cdots + a_n + \cdots .$$

To do this, one introduces the *partial sum* $S_n = \sum_{m=0}^{n} a_m \equiv a_0 + \cdots + a_n$ and takes $\sum_{m=0}^{\infty} a_m = \lim_{n\to\infty} S_n$ when the limit exists, in which case one says that the *series* $\sum_{m=0}^{\infty} a_m$ *converges*. Observe that if $\sum_{m=0}^{\infty} a_m$ converges in $\mathbb{R}$, then, by Cauchy's criterion, it must be true that $a_n = S_n - S_{n-1} \longrightarrow 0$ as $n \to \infty$.

The easiest series to deal with are those whose summands a_m are all non-negative. In this case the sequence $\{S_n : n \geq 0\}$ of partial sums is non-decreasing and therefore, depending on whether or not it stays bounded, it converges either to a finite number or to $+\infty$. This observation is useful even when the summands a_m's can take both signs. Namely, given $\{a_m : m \geq 0\}$, consider the series corresponding to $\{|a_m| : m \geq 0\}$. Clearly,

$$|S_{n_2} - S_{n_1}| = \left|\sum_{m=n_1+1}^{n_2} a_m\right| \leq \sum_{m=n_1+1}^{n_2} |a_m| = \sum_{m=0}^{\infty} |a_m| - \sum_{m=0}^{n_1} |a_m|$$

for $n_1 < n_2$. Hence, if $\sum_{m=0}^{\infty} |a_m| < \infty$ and therefore

$$\sum_{m=0}^{\infty} |a_m| - \sum_{m=0}^{n_1} |a_m| \longrightarrow 0 \quad \text{as } n_1 \to \infty,$$

then $\{S_n : n \geq 0\}$ satisfies Cauchy's criterion and is therefore convergent in $\mathbb{R}$. For this reason, one says that the series $\sum_{m=0}^{\infty} a_m$ is *absolutely convergent* if $\sum_{m=0}^{\infty} |a_m| < \infty$.

It is important to recognize that a series may be convergent in $\mathbb{R}$ even if it is not absolutely convergent.

Lemma 1.2.1 *Suppose $\{a_n : n \geq 0\}$ is a non-increasing sequence that converges to 0. Then the series $\sum_{m=0}^{\infty}(-1)^m a_m$ converges in $\mathbb{R}$. In fact,*

$$\left|\sum_{m=0}^{\infty}(-1)^m a_m - \sum_{m=0}^{n}(-1)^m a_m\right| \leq a_n.$$

Proof Set

$$T_n = \sum_{m=0}^{n}(-1)^m = \begin{cases} 1 & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

Then (cf. Exercise 1.6 for a generalization of this argument)

$$\begin{aligned}\sum_{m=0}^{n}(-1)^m a_m &= a_0 + \sum_{m=1}^{n}(T_m - T_{m-1})a_m = a_0 + \sum_{m=1}^{n} T_m a_m - \sum_{m=0}^{n-1} T_m a_{m+1} \\ &= T_n a_n + \sum_{m=0}^{n-1} T_m (a_m - a_{m+1}),\end{aligned}$$

and so

$$\sum_{m=0}^{n_2}(-1)^m a_m - \sum_{m=0}^{n_1}(-1)^m a_m = T_{n_2} a_{n_2} - T_{n_1} a_{n_1} + \sum_{m=n_1}^{n_2-1} T_m (a_m - a_{m+1}).$$

Note that

$$0 \le \sum_{m=n_1}^{n_2-1} T_m(a_m - a_{m+1}) \le \sum_{m=n_1}^{n_2-1} (a_m - a_{m+1}) = a_{n_1} - a_{n_2} \le a_{n_1},$$

and therefore $\left|\sum_{m=0}^{n_2}(-1)^m a_m - \sum_{m=0}^{n_1}(-1)^m a_m\right| \le 2a_{n_1}$, which, by Cauchy's criterion, shows that the series converges in $\mathbb{R}$. In addition, one has

$$-a_n \le -T_n a_n \le \sum_{m=0}^{\infty}(-1)^m a_m - \sum_{m=0}^{n}(-1)^m a_m \le a_n. \qquad \square$$

Lemma 1.2.1 provides lots of examples of series that are convergent but not absolutely convergent in $\mathbb{R}$. Indeed, although we know that $\lim_{m\to\infty} a_m = 0$ if $\sum_{m=0}^{\infty} a_m$ converges in $\mathbb{R}$, $\sum_{m=0}^{\infty} a_m$ need not converge in $\mathbb{R}$ just because $a_m \longrightarrow 0$. For example, since

$$\sum_{m=1}^{2^\ell-1} \frac{1}{m} = \sum_{k=0}^{\ell-1}\left(\sum_{m=2^k}^{2^{k+1}-1} \frac{1}{m}\right) \ge \frac{\ell}{2} \longrightarrow \infty \quad \text{as } \ell \to \infty,$$

the *harmonic series* $\sum_{m=1}^{\infty} \frac{1}{m}$ converges to ∞ and therefore does not converge in $\mathbb{R}$. Nonetheless, by Lemma 1.2.1, $\sum_{m=1}^{\infty} \frac{(-1)^n}{n}$ does converge in $\mathbb{R}$. In fact, see Exercise 1.6 to find out what it converges to.

A series that converges in $\mathbb{R}$ but is not absolutely convergent is said to be *conditionally convergent*. In general determining when a series is conditionally convergent can be very tricky (cf. Exercise 1.5 below), but there are many useful criteria for determining whether it is absolutely convergent. The basic reason is that if the series $\sum_{m=0}^{\infty} b_m$ is absolutely convergent and if $|a_m| \le C|b_m|$ for some $C < \infty$ and all $m \ge 0$, then $\sum_{m=0}^{\infty} a_m$ is also absolutely convergent. This simple observation is sometimes called the *comparison test*. One of its most frequent applications involves comparing the series under consideration to the *geometric series* $\sum_{n=0}^{\infty} r^m$, where $0 < r < 1$. If $S_n = \sum_{m=0}^{n} r^m$, then $rS_n = \sum_{m=1}^{n+1} r^m = S_n + r^{n+1} - 1$, and therefore $S_n = \frac{1-r^{n+1}}{1-r}$, which shows that $\sum_{m=0}^{\infty} r^m = \frac{1}{1-r} < \infty$. One reason why the geometric series arises in applications of the comparison test is the following. Suppose that $\{a_m : m \ge 0\} \subseteq \mathbb{R}$ and that $|a_{m+1}| \le r|a_m|$ for some $r \in (0,1)$ and all $m \ge m_0$. Then $|a_m| \le |a_{m_0}|r^{m-m_0}$ for all $m \ge m_0$, and so, if $C = r^{-m_0}\max\{|a_0|, \dots, |a_{m_0}|\}$, then $|a_m| \le Cr^m$ for all $m \ge 0$. For obvious reasons, the resulting criterion is called the *ratio test*.

Series whose summands decay at least as fast as those of the geometric series are considered to be converging very fast and by no means include all absolutely convergent series. For example, let $\alpha > 1$, and consider the series $\sum_{m=1}^{\infty} \frac{1}{m^\alpha}$. To see that this series is absolutely convergent, we use the same idea as we did when we showed that the harmonic series diverges. That is,

$$\sum_{m=1}^{2^\ell-1}\frac{1}{m^\alpha}=\sum_{k=0}^{\ell-1}\left(\sum_{m=2^k}^{2^{k+1}-1}\frac{1}{m^\alpha}\right)\le\sum_{k=0}^{\ell-1}2^{(1-\alpha)k}\le\frac{1}{1-2^{1-\alpha}}<\infty.$$

Of course, if $\alpha\in[0,1]$, then $n^{-\alpha}\ge n^{-1}$, and therefore $\sum_{n=1}^\infty\frac{1}{n^\alpha}=\infty$ since $\sum_{n=1}^\infty\frac{1}{n}=\infty$.

1.3 Topology of and Continuous Functions on $\mathbb{R}$

A subset I of $\mathbb{R}$ is called an *interval* if $x\in I$ whenever there exist $a,\ b\in I$ for which $a<x<b$. In particular, if $a,\ b\in\mathbb{R}$ with $a\le b$, then

$$\begin{aligned}&(a,b)=\{x:\ a<x<b\},\quad (a,b]=\{x:\ a<x\le b\},\\&[a,b)=\{x:\ a\le x<b\},\quad [a,b]=\{x:\ a\le x\le b\},\\&(-\infty,a]=\{x:\ x\le a\},\quad (-\infty,a)=\{x:\ x<a\},\\&[b,\infty)=\{x:\ x\ge b\},\quad \text{and } (b,\infty)=\{x:\ x>b\}\end{aligned}\tag{1.3.1}$$

are all intervals, as is $\mathbb{R}=(-\infty,\infty)$. Note that if $a=b$, then (a,b), $(a,b]$, and $[a,b)$ are all empty. Thus, the empty set $\emptyset$ is an interval.

A subset $G\subseteq\mathbb{R}$ is said to be *open* if either $G=\emptyset$ or, for each $x\in G$, there is a $\delta>0$ such that $(x-\delta,x+\delta)\subseteq G$. Clearly, the union of any number of open sets is again open, and the intersection of a finite number of open sets is again open. However, the countable intersection of open sets may not be open. For example, the interval (a,b) is open for every $a<b$, but

$$\{0\}=\bigcap_{n=1}^\infty\left(-\tfrac{1}{n},\tfrac{1}{n}\right)$$

is not open.

A subset $F\subseteq\mathbb{R}$ is said to be *closed* if its *complement* $F\complement=\mathbb{R}\setminus F$ is open. Thus, the intersection of an arbitrary number of closed set is closed, as is the union of a finite number of them. Given any set $S\subseteq\mathbb{R}$, the *interior* $\operatorname{int}(S)$ of S is the largest open subset of S. Equivalently, it is the union of all the open subsets of S. Similarly, the *closure* $\bar{S}$ of S is the smallest closed set containing S, or, equivalently, it is the intersection of all the closed sets containing S. Clearly, an open set is equal to its own interior, and any closed set is equal to its own closure.

Lemma 1.3.1 *The subset F is closed if and only if $x\in F$ whenever there exists a sequence $\{x_n:\ n\ge 1\}\subseteq F$ such that $x_n\to x$. Moreover, for any $S\subseteq\mathbb{R}$, $x\in\operatorname{int}(S)$ if and only if $(x-\delta,x+\delta)\subseteq S$ for some $\delta>0$, and $x\in\bar{S}$ if and only if there is a sequence $\{x_n:\ n\ge 1\}\subseteq S$ such that $x_n\to x$.*

Proof Suppose that F is closed, and set $G = F\complement$. Let $\{x_n : n \geq 1\} \subseteq F$ be a sequence that converges to x. If x were not in F, it would be an element of the open set G. Thus there would exist a $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq G$. But, because $x_n \to x$, $|x_n - x| < \delta$ and therefore $x_n \in (x - \delta, x + \delta)$ for all sufficiently large n's, and this would mean that $x_n \in F \cap G$ for all large enough n's. Since $F \cap G$ is empty, this shows that x must have been in F.

Now assume that $x \in F$ whenever there exists a sequence $\{x_n : n \geq 1\} \subseteq F$ such that $x_n \to x$, and set $G = F\complement$. To see that G is open, suppose not. Then there would be some $x \in G$ with the property that for all $n \geq 1$ there is an $x_n \in F$ such that $|x - x_n| < \frac{1}{n}$. But this would mean that $x_n \to x$ and therefore that $x \in F$.

Suppose that $x \in \text{int}(S)$. Then there is an open G for which $x \in G \subseteq S$. Hence there is a $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq G \subseteq S$. Conversely, if there is a $\delta > 0$ for which $(x - \delta, x + \delta) \subseteq S$, then, since $(x - \delta, x + \delta)$ is an open subset of S, $x \in \text{int}(S)$.

Finally, let x be an element of $\mathbb{R}$ to which no sequence $\{x_n : n \geq 1\} \subseteq S$ converges. Then there must exist a $\delta > 0$ such that $(x - \delta, x + \delta) \cap S = \emptyset$. To see this, suppose that no such $\delta > 0$ exists. Then, for each $n \geq 1$, there would exist an $x_n \in S$ such that $|x - x_n| < \frac{1}{n}$, which would mean that $x_n \to x$. Hence such a $\delta > 0$ exists. But if $(x - \delta, x + \delta) \cap S = \emptyset$, then S is contained in the closed set $(x - \delta, x + \delta)\complement$, and therefore $x \notin \bar{S}$. Conversely, if $x_n \to x$ for some sequence $\{x_n : n \geq 1\} \subseteq S$, then $x \in F$ for every closed $F \supseteq S$, and therefore $x \in \bar{S}$. $\square$

It should be clear that (a, b) and $[a, b]$ are, respectively, the interior and closure of $[a, b]$, $(a, b]$, $[a, b)$, and (a, b), that (a, ∞) and $[a, \infty)$ are, respectively, the interior and closure of $[a, \infty)$ and (a, ∞), and that $(-\infty, b)$ and $(-\infty, b]$ are, respectively, the interior and closure of $(-\infty, b)$ and $(-\infty, b]$. Finally, $\mathbb{R} = (-\infty, \infty)$ and $\emptyset$ are both open and closed.

Lemma 1.3.2 *Given a non-empty $S \subseteq \mathbb{R}$, set $I = \{y : y \geq x$ for all $x \in S\}$. Then I is a closed interval. Moreover, if S is bounded above (i.e., there is an $M < \infty$ such that $x \leq M$ for all $x \in S$), then there is a unique element $\sup S \in \mathbb{R}$ with the property that $y \geq \sup S \geq x$ for all $x \in S$ and $y \in I$. Equivalently, $I = \big[\sup S, \infty\big)$.*

Proof It is clear that I is a closed interval that is non-empty if S is bounded above. In addition, if both y_1 and y_2 have the property specified for $\sup S$, then $y_1 \leq y_2$ and $y_2 \leq y_1$, and therefore $y_1 = y_2$.

Assuming that S is bounded above, we will now show that, for any $\delta > 0$, there exist an $x \in S$ and a $y \in I$ such that $y - x < \delta$. Indeed, suppose this were not true. Then there would exist a $\delta > 0$ such that $y \geq x + 2\delta$ for all $x \in S$ and $y \in I$. Set $S' = \{x + \delta : x \in S\}$. Then $S' \cap I = \emptyset$, and so for every $x \in S$ there would exist an $x' \in S$ such that $x' \geq x + \delta$. But this would mean that for any $x \in S$ and $n \geq 1$, there would exist an $x_n \in S$ such that $x_n \geq x + n\delta$, which, because S is bounded above, is impossible.

For each $n \geq 1$, choose $x_n \in S$ and $\eta_n \in I$ so that $\eta_n \leq x_n + \frac{1}{n}$, and set $y_n = \eta_1 \wedge \cdots \wedge \eta_n$. Then $\{y_n : n \geq 1\}$ is a non-increasing sequence in I, and so it has a limit $c \in I$. Obviously, $[c, \infty) \subseteq I$, and therefore $x \leq c$ for all $x \in S$. Finally, suppose that $y \in I$. Then $y + \frac{1}{n} \geq x_n + \frac{1}{n} \geq y_n \geq c$ for all $n \geq 1$, and so $y \geq c$. Hence, $c = \sup S$ and $I = \big[\sup_I S, \infty)$. □

The number $\sup S$ is called the *supremum* of S. If $\emptyset \neq S \subseteq \mathbb{R}$ is bounded below, then there is a unique element $\inf S \in \mathbb{R}$, known as the *infimum* of S, such that $x \geq \inf S$ for all $x \in S$ and $y \leq \inf S$ if $y \leq x$ for all $x \in S$. To see this, simply take $\inf S = -\sup\{-x : x \in S\}$. Starting from Lemma 1.3.2, it is an easy matter to show that any non-empty interval I is one of the sets in (1.3.1), where $a = \inf I$ if I is bounded below and $b = \sup I$ if I is bounded above. Notice that, although one or both $\sup S$ and $\inf S$ may not be elements of S, whenever one of them exists, it must be an element of $\bar{S}$. When $\sup S \in S$, it is often referred to as the *maximum* of S and is written as $\max S$ instead of $\sup S$. Similarly, when $\inf S \in S$, it is called the *minimum* of S and is denoted by $\min S$. Finally, we will take $\sup S = \infty$ if S is unbounded above and $\inf S = -\infty$ if it is unbounded below. Notice that if $\{x_n : n \geq 1\}$ is non-decreasing, then $x_n \longrightarrow \sup_{n\geq 1} x_n \equiv \sup\{x_n : n \geq 1\}$, and that $x_n \longrightarrow \inf_{n\geq 1} x_n$ if it is non-increasing.

Given a bounded sequence $\{x_n : n \geq 1\} \subseteq \mathbb{R}$, set $y_m = \inf\{x_n : n \geq m\}$ and $z_m = \sup\{x_n : n \geq m\}$. Then $\{y_m : m \geq 1\}$ is a non-decreasing sequence that is bounded above, and $\{z_m : m \geq 1\}$ is a non-increasing sequence that is bounded below. Thus the *limit inferior* and *limit superior*, respectively,

$$\varliminf_{n\to\infty} x_n \equiv \lim_{m\to\infty} \inf\{x_n : n \geq m\} \quad \text{and} \quad \varlimsup_{n\to\infty} x_n \equiv \lim_{m\to\infty} \sup\{x_n : n \geq m\}$$

exist in $\mathbb{R}$. If $\{x_n : n \geq 1\}$ is not bounded above, then $\varlimsup_{n\to\infty} x_n \equiv \infty$, and if it is unbounded below, then $\varliminf_{n\to\infty} x_n \equiv -\infty$.

Theorem 1.3.3 *Let $\{x_n : n \geq 1\}$ be a bounded sequence. Then both $\varliminf_{n\to\infty} x_n$ and $\varlimsup_{n\to\infty} x_n$ are limit points of $\{x_n : n \geq 1\}$. Furthermore, every limit point lies in between these two.*

Proof Set $z_m = \sup\{x_n : n \geq m\}$. Then $z_m \longrightarrow z = \varlimsup_{n\to\infty} x_n$. Take $m_0 = 0$, and, given $m_0, \dots, m_{k-1}$, take $m_k > m_{k-1}$ so that $|x_{m_k} - z_{m_{k-1}+1}| \leq \frac{1}{k}$. Then, since $z_{m_{k-1}+1} \longrightarrow z$,

$$|x_{m_k} - z| \leq |x_{m_k} - z_{m_{k-1}+1}| + |z_{m_{k-1}+1} - z| \leq \frac{1}{k} + |z_{m_{k-1}+1} - z| \longrightarrow 0$$

as $k \to \infty$. After replacing $\{x_n : n \geq 1\}$ by $\{-x_n : n \geq 1\}$, one gets the same conclusion about $\varliminf_{n\to\infty} x_n$. Finally, if $\{x_{n_k} : k \geq 1\}$ is any subsequence that converges to some x, then $x_{n_k} \leq z_{n_k}$ and so $x \leq \varlimsup_{n\to\infty} x_n$. Similarly, $x \geq \varliminf_{n\to\infty} x_n$. □

A non-empty open set is said to be *connected* if and only if it cannot be written as the union of two, disjoint, non-empty open sets. See Exercise 4.5 for more information.

Lemma 1.3.4 *Let G be a non-empty open set. Then G is connected if and only it is an open interval.*

Proof Suppose that G is connected. If G were not an interval, then there would exist $a, b \in G$ and $c \notin G$ such that $a < c < b$. But then $a \in G_1 \equiv G \cap (-\infty, c)$, $b \in G_2 \equiv G \cap (c, \infty)$, G_1 and G_2 are open, $G_1 \cap G_2 = \emptyset$, and $G = G_1 \cup G_2$. Thus G must be an interval.

Now assume that G is a non-empty open interval, and suppose that $G = G_1 \cup G_2$, where G_1 and G_2 are disjoint open sets. If G_1 and G_2 were non-empty, then we could find $a \in G_1$ and $b \in G_2$, and, without loss in generality, we could assume that $a < b$. Set $c = \inf\{x \in (a, b) : x \notin G_1\}$. Because $[a, b] \subseteq G$, $c \in G$, and because $G_1 \complement$ is closed, $c \notin G_1$. Hence $c \in G_2$. But G_2 is open, and therefore there exists an $x \in G_2 \cap (a, c)$, which would mean that $c \neq \inf\{x \in (a, b) : x \notin G_1\}$. Thus, either G_1 or G_2 must have been empty. □

If $\emptyset \neq S \subseteq \mathbb{R}$ and $f : S \longrightarrow \mathbb{R}$ (i.e., f is an $\mathbb{R}$-valued function on S), then f is said to be *continuous at* an $x \in S$ if, for each $\epsilon > 0$, there is a $\delta > 0$ such that $|f(x') - f(x)| < \epsilon$ whenever $x' \in S$ and $|x' - x| < \delta$. A function is said to be *continuous on* S if it is continuous at each $x \in S$. For example, the function $f : \mathbb{R} \longrightarrow [0, \infty)$ given by $f(x) = |x|$ is continuous because, by the triangle inequality, $|x| \leq |y - x| + |y|$, $|y| \leq |y - x| + |x|$, and therefore $\big||y| - |x|\big| \leq |y - x|$.

Lemma 1.3.5 *Let H be a non-empty subset of $\mathbb{R}$ and $f : H \longrightarrow \mathbb{R}$. Then f is continuous at $x \in H$ if and only if $f(x_n) \longrightarrow f(x)$ for every sequence $\{x_n : n \geq 1\} \subseteq H$ with $x_n \to x$. Moreover, if H is open, then f is continuous on H if and only if*

$$f^{-1}(G) \equiv \{x \in H : f(x) \in G\} \quad \textit{is open for every open } G \subseteq \mathbb{R}.$$

Proof If f is continuous at $x \in H$ and $\epsilon > 0$, choose $\delta > 0$ so that $|f(x') - f(x)| < \epsilon$ when $x' \in H \cap (x - \delta, x + \delta)$. Given $\{x_n : n \geq 1\} \subseteq H$ with $x_n \longrightarrow x$, choose n_ϵ so that $|x_n - x| < \delta$ when $n \geq n_\epsilon$. Then $|f(x_n) - f(x)| < \epsilon$ for all $n \geq n_\epsilon$. Conversely, suppose that $f(x_n) \longrightarrow f(x)$ whenever $\{x_n : n \geq 1\} \subseteq H$ converges to x. If f were not continuous at x, then there would exist an $\epsilon > 0$ such that, for each $n \geq 1$ there is an $x_n \in H \cap \big(x - \frac{1}{n}, x + \frac{1}{n}\big)$ for which $|f(x_n) - f(x)| \geq \epsilon$. But, since $x_n \to x$, no such sequence can exist.

Suppose that H is open and that f is continuous on H, and let G be open. Given $x \in f^{-1}(G)$, set $y = f(x)$. Then, because G is open, there exists an $\epsilon > 0$ such that $(y - \epsilon, y + \epsilon) \subseteq G$, and because f is continuous and H is open, there is a $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq H$ and $|f(\xi) - y| < \epsilon$ whenever $\xi \in (x - \delta, x + \delta)$. Hence $(x - \delta, x + \delta) \subseteq f^{-1}(G)$. Conversely, assume that $f^{-1}(G)$ is open whenever G is. To see that f must be continuous, suppose that $x \in H$, and set $y = f(x)$. Given $\epsilon > 0$, x is an element of the open set $f^{-1}\big((y - \epsilon, y + \epsilon)\big)$, and so there exists a $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq f^{-1}\big((y - \epsilon, y + \epsilon)\big)$. Hence $|f(\xi) - f(x)| < \epsilon$ whenever $|\xi - x| < \delta$. □

From the first part of Lemma 1.3.5 combined with the properties of convergent sequences discussed in Sect. 1.1, it follows that linear combinations, products, and, as long as the denominator doesn't vanish, quotients of continuous functions are again continuous. In particular, it is easy to check that polynomials are continuous.

Theorem 1.3.6 (Intermediate Value Theorem) *Suppose that* $-\infty < a < b < \infty$ *and that* $f : [a,b] \longrightarrow \mathbb{R}$ *is continuous. Then for every*

$$y \in \big(f(a) \wedge f(b), f(a) \vee f(b)\big)$$

there is an $x \in (a,b)$ *such that* $y = f(x)$. *In particular, the interior of the image of a open interval under a continuous function is again an open interval.*

Proof There is nothing to do if $f(a) = f(b)$, and so, without loss in generality, assume that $f(a) < f(b)$. Given $y \in \big(f(a), f(b)\big)$, set

$$G_1 = G \cap \{x \in (a,b) : \ f(x) < y\} \text{ and } G_2 = G \cap \{x \in (a,b) : \ f(x) > y\}.$$

Then G_1 and G_2 are disjoint, non-empty subsets of sets (a,b) which, by Lemma 1.3.5, are open. Hence, by Lemma 1.3.4, there exists an $x \in (a,b) \setminus (G_1 \cup G_2)$, and clearly $y = f(x)$. □

Here is a second, less geometric, proof of this theorem. Assume that $f(a) < f(b)$ and that $y \in \big(f(a), f(b)\big)$ is not in the image of (a,b) under f. Define $H = \{x \in [a,b] : \ f(x) < y\}$, and let $c = \sup H$. By continuity, $f(c) \le y$. Suppose $\epsilon \equiv y - f(c) > 0$. Choose $0 < \delta < \frac{b-c}{2}$ so that $|f(x) - f(c)| < \epsilon$ whenever $0 < x - c < 2\delta$. Then $f(c+\delta) = f(c) + \big(f(c+\delta) - f(c)\big) < f(c) + y - f(c) = y$, and so $c + \delta \in H$. But this leads to the contradiction $c + \delta \le \sup H = c$.

A function that has the property proved for continuous functions in this theorem is said to have the *intermediate value property*, and, at one time, it was used as the defining property for continuity because it is means that the graph of the function can be drawn without "lifting the pencil from the page". However, although it is implied by continuity, it does not imply continuity. For example, consider the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by $f(x) = \sin \frac{1}{x}$ when $x \neq 0$ and $f(0) = 0$. Obviously, f is discontinuous at 0, but it nonetheless has the intermediate value property. See Exercise 1.15 for a more general source of discontinuous functions that have the intermediate value property.

Here is an interesting corollary.

Corollary 1.3.7 *Let* $a < b$ *and* $c < d$, *and suppose that* f *is a continuous, one-to-one mapping that takes* $[a,b]$ *onto* $[c,d]$. *Then either* $f(a) = c$ *and* f *is strictly increasing (i.e.,* $f(x) < f(y)$ *if* $x < y$*), or* $f(a) = d$ *and* f *is strictly decreasing.*

Proof We first show that $f(a)$ must equal either c or d. To this end, suppose that $c = f(s)$ for some $s \in (a,b)$, and choose $0 < \delta < (b-s) \wedge (s-a)$. Then both $f(s-\delta)$ and $f(s+\delta)$ lie in $(c,d]$. Choose $c < t < f(s-\delta) \wedge f(s+\delta)$. Then, by Theorem 1.3.6, there exist $s_1 \in (s-\delta, s)$ and $s_2 \in (s, s+\delta)$, such that $f(s_1) = t = f(s_2)$. But this would mean that f is not one-to-one, and therefore we know that $f(s) \neq c$ for any $s \in (a, b)$. Similarly, $f(s) \neq d$ for any $s \in (a, b)$. Hence, either $f(a) = c$ and $f(b) = d$ or $f(a) = d$ and $f(b) = c$.

Now assume that $f(a) = c$ and $f(b) = d$. If f were not increasing, then there would be $a < s_1 < s_2 < b$ such that $c < t_2 = f(s_2) < f(s_1) = t_1 < d$. Choose $t_2 < t < t_1$. Then, by Theorem 1.3.6, there would exist an $s_3 \in (a, s_1)$ and an $s_4 \in (s_1, s_2)$ such that $f(s_3) = t = f(s_4)$, which would mean that f could not be one-to-one. Hence f is non-decreasing, and because it is one-to-one, it must be strictly increasing. Similarly, if $f(a) = d$ and $f(b) = c$, then f must be strictly decreasing. □

1.4 More Properties of Continuous Functions

A real-valued function f on a non-empty set $S \subseteq \mathbb{R}$ is said to be *uniformly continuous* on S if for each $\epsilon > 0$ there is a $\delta > 0$ such that $|f(x') - f(x)| < \epsilon$ for all $x,\ x' \in S$ with $|x' - x| < \delta$. In other words, the choice of δ depends only on ϵ and not on the point under consideration.

Theorem 1.4.1 *Let K be a bounded, closed, non-empty subset of $\mathbb{R}$. If $f : K \longrightarrow \mathbb{R}$ is continuous on K, then it is bounded and uniformly continuous there. Moreover, there is at least one point in K at which f achieves its maximum value, and at least one at which it achieves its minimum value. Hence, if K is a bounded, closed interval, then f takes every value between its maximum and minimum values.*

Proof Suppose f were not uniformly continuous. Then, for some $\epsilon > 0$, there would exist sequences $\{x_n : n \geq 1\} \subseteq K$ and $\{x'_n : n \geq 1\} \subseteq K$ such that $|x'_n - x_n| < \frac{1}{n}$ and yet $|f(x'_n) - f(x_n)| \geq \epsilon$. Moreover, by Theorem 1.3.3, we could and will assume that $x_n \to x \in K$, in which case $x'_n \to x$ as well. Now, set $x''_{2n-1} = x_n$ and $x''_{2n} = x'_n$ for $n \geq 1$. Then $x''_n \to x$ and $|f(x''_{2n}) - f(x''_{2n-1})| \geq \epsilon$, which is impossible since f is continuous at x.

To see that f is bounded, suppose it is not. Then for each $n \geq 1$ there exists an $x_n \in K$ at which $|f(x_n)| \geq n$. Choose a subsequence $\{x_{n_k} : k \geq 1\}$ that converges to a point $x \in K$. Then, because (cf. Exercise 1.9) $|f|$ is continuous,

$$\infty > |f(x)| = \lim_{k\to\infty} |f(x_{n_k})| \geq \lim_{k\to\infty} n_k = \infty,$$

which is impossible. Knowing that f is bounded, set $M = \sup\{f(x) : x \in K\}$, and choose $\{x_n : n \geq 1\} \subseteq K$ so that $f(x_n) \longrightarrow M$ as $n \to \infty$. Now, just as before, choose a subsequence that converges to a point x, and conclude that $f(x) = M$. The same argument shows that f achieve its minimum value. Finally, when K is an interval, Theorem 1.3.6 implies the concluding assertion. □

It is often important to know when a function defined on a set admits a continuous extension to the closure. With this in mind, say a subset D of a set S is *dense* in S if $\bar{D} \supseteq S$.

Lemma 1.4.2 *Suppose that D is a dense subset of a non-empty closed set F and that $f : D \longrightarrow \mathbb{R}$ is uniformly continuous. Then there is a unique extension $\bar{f} : F \longrightarrow \mathbb{R}$ of f as a continuous function, and the extension is uniformly continuous on F.*

Proof The uniqueness is obvious. Indeed, if $\bar{f}_1$ and $\bar{f}_2$ were two continuous extensions and $x \in F$, choose $\{x_n : n \geq 1\} \subseteq D$ so that $x_n \to x$ and therefore $\bar{f}_1(x) = \lim_{n\to\infty} f(x_n) = \bar{f}_2(x)$.

For each $\epsilon > 0$ choose $\delta_\epsilon > 0$ such that $|f(x') - f(x)| < \epsilon$ for all $x, x' \in D$ with $|x' - x| < \delta_\epsilon$. Given $x \in F$, choose $\{x_n : n \geq 1\} \subseteq D$ so the $x_n \to x$, and, for $\epsilon > 0$, choose n_ϵ so that $|x_n - x| < \delta_{\frac{\epsilon}{2}}$ if $n \geq n_\epsilon$. Then,

$$|f(x_n) - f(x_m)| \leq |f(x_n) - f(x)| + |f(x) - f(x_m)| < \epsilon \quad \text{for all } m, n \geq n_\epsilon.$$

Hence, by Cauchy's criterion, $\{f(x_n) : n \geq 1\}$ is convergent in $\mathbb{R}$. Furthermore, the limit does not depend on the sequence chosen. In fact, given an $x' \in F$ for which $|x' - x| < \frac{\delta_\epsilon}{3}$ and a sequence $\{x'_n : n \geq 1\} \subseteq D$ that converges to x', choose n'_ϵ so that $|x_n - x| \vee |x'_n - x'| < \frac{\delta_\epsilon}{3}$ for $n \geq n'_\epsilon$. Then

$$|x'_n - x_n| \leq |x_n - x| + |x - x'| + |x' - x'_n| < \delta_\epsilon$$

and therefore $|f(x'_n) - f(x_n)| < \epsilon$ for $n \geq n'_\epsilon$. Thus

$$\left|\lim_{n\to\infty} f(x'_n) - \lim_{n\to\infty} f(x_n)\right| \leq \epsilon.$$

In particular, if $x' = x$, then $\lim_{n\to\infty} f(x'_n) = \lim_{n\to\infty} f(x_n)$, and so we can unambiguously define $\bar{f}(x) = \lim_{n\to\infty} f(x_n)$ using any $\{x_n : n \geq 1\} \subseteq D$ with $x_n \to x$. Moreover, if $x' \in F$ and $|x' - x| < \frac{\delta_\epsilon}{3}$, then $|\bar{f}(x') - \bar{f}(x)| \leq \epsilon$, and so $\bar{f}$ is uniformly continuous on F. □

The next theorem deals with the possibility of taking the *inverse* of functions on $\mathbb{R}$.

Theorem 1.4.3 *Let $\emptyset \neq S \subseteq \mathbb{R}$, and suppose $f : S \longrightarrow \mathbb{R}$ is strictly increasing. Then, for each $y \in \mathbb{R}$, there is at most one $x \in S$ such that $y = f(x)$. Next assume that $f : [a, b] \longrightarrow \mathbb{R}$ is strictly increasing and continuous. Then, there is a unique function $f^{-1} : \big[f(a), f(b)\big] \longrightarrow [a, b]$ such that $f\big(f^{-1}(y)\big) = y$ for $y \in \big[f(a), f(b)\big]$. Moreover, $f^{-1}\big(f(x)\big) = x$ for $x \in [a, b]$, and f^{-1} is continuous and strictly increasing.*

Proof Since $x_1 < x_2 \implies f(x_1) < f(x_2)$, it is clear that there is at most one x such that $y = f(x)$. Now assume that $S = [a, b]$ and f is continuous. By Theorem 1.3.6 we know that for each $y \in \big[f(a), f(b)\big]$ there is a necessarily unique $x \in [a, b]$ such that $y = f(x)$. Thus, we can define a function $f^{-1} : \big[f(a), f(b)\big] \longrightarrow [a, b]$ so that $f^{-1}(y)$ is the unique $x \in [a, b]$ for which $y = f(x)$. To see that $f^{-1}\big(f(x)\big) = x$ for $x \in [a, b]$, set $y = f(x)$ and $x' = f^{-1}(y)$. Then $f(x') = y = f(x)$, and so $x' = x$. Finally, to show that f^{-1} is continuous, suppose that $\{y_n : n \geq 1\} \subseteq \big[f(a), f(b)\big]$

and that $y_n \to y$. Set $x_n = f^{-1}(y_n)$ and $x = f^{-1}(y)$. If $x_n \not\to x$, then there would exist an $\epsilon > 0$ and a subsequence $\{x_{n_k} : k \geq 1\}$ such that $|x_{n_k} - x| \geq \epsilon$. Further, by Theorem 1.3.3, we could choose this subsequence to be convergent to some point x'. But this would mean that $|x' - x| \geq \epsilon$ and yet $f(x') = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} y_{n_k} = f(x)$, which is impossible. □

We conclude this discussion about continuous functions with a result that shows that continuity is preserved under a certain type of limit. Specifically, say that a sequence $\{f_n : n \geq 1\}$ on S *converges uniformly* on S to a function f on S if $\lim_{n\to\infty} \sup_{x\in S} |f_n(x) - f(x)| = 0$.

Lemma 1.4.4 *Suppose that* $\{f_n : n \geq 1\}$ *is a sequence of continuous functions on a set* S*. If* $\{f_n : n \geq 1\}$ *converges uniformly on* S *to a function* f*, then* f *is continuous. Furthermore, if*

$$\lim_{m\to\infty} \sup_{n>m} \sup_{x\in S} |f_n(x) - f_m(x)| = 0,$$

then there is an f *on* S *to which* $\{f_n : n \geq 1\}$ *converges uniformly.*

Proof Given $\epsilon > 0$ and $x \in S$, choose n so that $|f_n(y) - f(y)| < \frac{\epsilon}{3}$ for all $y \in S$, and choose $\delta > 0$ so that $|f_n(y) - f_n(x)| < \frac{\epsilon}{3}$ for $y \in S \cap (x - \delta, x + \delta)$. Then

$$|f(y) - f(x)| \leq |f(y) - f_n(y)| + |f_n(y) - f_n(x)| + |f_n(x) - f(x)| < \epsilon$$

for $y \in S \cap (x - \delta, x + \delta)$, which means that f is continuous.

To prove the second part, first note that, by Cauchy's convergence criterion, for each $x \in S$ there is an $f(x)$ to which $\{f_n(x) : n \geq 1\}$ converges. Given $\epsilon > 0$, choose m_ϵ so that $\sup_{x\in S} |f_n(x) - f_m(x)| < \epsilon$ if $n > m \geq m_\epsilon$. Then, for $x \in S$ and $m \geq m_\epsilon$,

$$|f(x) - f_m(x)| = \lim_{n\to\infty} |f_n(x) - f_m(x)| \leq \epsilon.$$

□

1.5 Differentiable Functions

A function f on an open set G is said to be *differentiable* at a point $x \in G$ if the limit

$$f'(x) = \partial f(x) = \frac{df}{dx}(x) \equiv \lim_{y\to x} \frac{f(y) - f(x)}{y - x}$$

exists in $\mathbb{R}$. That is, for all $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\left| \frac{f(y) - f(x)}{y - x} - f'(x) \right| < \epsilon \quad \text{if } y \in G \text{ and } 0 < |y - x| < \delta.$$

Clearly, what $f'(x)$ represents is the instantaneous rate at which f is changing at x. More geometrically, if one thinks in terms of the graph of f, then $\frac{f(y)-f(x)}{y-x}$ is the *slope* of the line connecting $\big(x, f(x)\big)$ to $\big(y, f(y)\big)$, and so, when $y \to x$, this ratio should represent the slope of the tangent line to the graph at $\big(x, f(x)\big)$. Alternatively, thinking of x as time and $f(x)$ as the distance traveled up to time x, $f'(x)$ is the instantaneous velocity at time x.

When $f'(x)$ exists, it is called the *derivative* of f at x. Obviously, if f is differentiable at x, then it is continuous there, and, in fact, the rate at which $f(y)$ approaches $f(x)$ is commensurate with $|y-x|$. However, just because $|f(y)-f(x)| \le C|y-x|$ for some $C < \infty$ does not guaranty that f is differentiable at x. For example, if $f(x) = |x|$, then $\big|f(y) - f(x)\big| \le |y - x|$ for all $x, y \in \mathbb{R}$, but f is not differentiable at 0, since $\frac{f(y)-f(0)}{y-0}$ is 1 when $y > 0$ and -1 when $y < 0$.

When f is differentiable at every $x \in G$ it is said to be *differentiable on* G, and when it is differentiable on G and its derivative f' is continuous there, f is said to be *continuously differentiable*.

Differentiability is preserved under the basic arithmetic operations. To be precise, assume that f and g are differentiable at x. If $\alpha, \beta \in \mathbb{R}$, then

$$\begin{aligned}&\left|\frac{\big(\alpha f(y) + \beta f(y)\big) - \big(\alpha f(x) + \beta g(x)\big)}{y-x} - \big(\alpha f'(x) + \beta g'(y)\big)\right| \\ &\le |\alpha|\left|\frac{f(y)-f(x)}{y-x} - f'(x)\right| + |\beta|\left|\frac{g(y)-g(x)}{y-x} - g'(x)\right|,\end{aligned}$$

and so $(\alpha f + \beta g)'(x)$ exists and is equal to $\alpha f'(x) + \beta g'(x)$. More interesting, by adding and subtracting $f(x)g(y)$ in the numerator, one sees that

$$\begin{aligned}&\left|\frac{f(y)g(y) - f(x)g(x)}{y-x} - \big(f'(x)g(x) + f(x)g'(x)\big)\right| \\ &\le \left|\frac{\big(f(y)-f(x)\big)g(y)}{y-x} - f'(x)g(x)\right| + \left|\frac{f(x)\big(g(y)-g(x)\big)}{y-x} - f(x)g'(x)\right|,\end{aligned}$$

and therefore fg is differentiable at x and $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$. This important fact is called the *product rule* or sometimes *Leibniz's formula*. Finally, if $g(x) \ne 0$, then, since g is continuous at x, $g(y) \ne 0$ for y sufficiently near x, and

$$\begin{aligned}&\left|\frac{\frac{f}{g}(y) - \frac{f}{g}(x)}{y-x} - \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}\right| \\ &\le \left|\frac{f(y)-f(x)}{(y-x)g(y)} - \frac{f'(x)}{g(x)}\right| + \left|\frac{f(x)\big(g(y)-g(x)\big)}{(y-x)g(x)g(y)} - \frac{f(x)g'(x)}{g(x)^2}\right|,\end{aligned}$$

which means that $\frac{f}{g}$ is differentiable at x and that $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$. This is often called the *quotient rule*.

We now have enough machinery to show that lots of functions are continuously differentiable. To begin with, consider the function $f_k(x) = x^k$ for $k \in \mathbb{N}$. Clearly, $f_0' = 0$ and $f_1' = 1$. Using the product rule and induction on $k \geq 1$, one can show that $f_k'(x) = kx^{k-1}$. Indeed, if this is true for k, then, since $f_{k+1} = f_1 f_k$, $f_{k+1}'(x) = f_k(x) + xf_k'(x) = (k+1)x^k$. Alternatively, one can use the identity $y^k - x^k = (y-x)\sum_{m=1}^{k} x^m y^{k-m-1}$ to see that

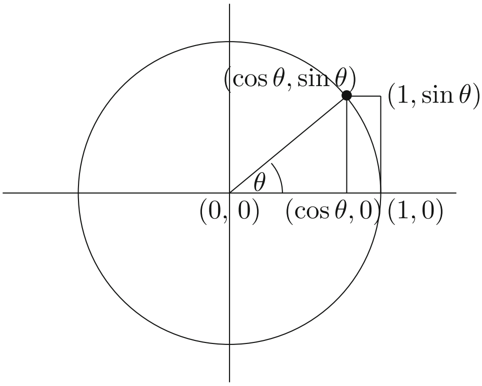
$$\left|\frac{y^k - x^k}{y-x} - kx^{k-1}\right| \leq \left|\sum_{m=1}^{k}(x^m y^{k-m-1} - x^{k-1})\right|.$$

As a consequence, we now see that polynomials $\sum_{m=0}^{n} a_m x^m$ are continuously differentiable on $\mathbb{R}$ and that

$$\left(\sum_{m=0}^{n} a_m x^m\right)' = \sum_{m=1}^{n} m a_m x^{m-1}.$$

In addition, if $k \geq 1$ and $x \neq 0$, then, by the quotient rule, $(x^{-k})' = \left(\frac{1}{x^k}\right)' = -\frac{kx^{k-1}}{x^{2k}} = -kx^{-k-1}$. Hence for any $k \in \mathbb{Z}$, $(x^k)' = kx^{k-1}$ with the understanding that $x \neq 0$ when $k < 0$.

A more challenging source of examples are the trigonometric functions sine and cosine. For our purposes, sin and cos should be thought about in terms of the unit circle (i.e., the perimeter of the disk of radius 1 centered at the origin) $\mathbb{S}^1(\mathbf{0}, 1)$ in $\mathbb{R}^2$. That is, if one starts at $(1, 0)$ and travels counterclockwise along $\mathbb{S}^1(\mathbf{0}, 1)$ for a distance $\theta \geq 0$, then $(\cos\theta, \sin\theta)$ is the Cartesian representation of the point at which one arrives. Similarly, if $\theta < 0$, then $(\cos\theta, \sin\theta)$ is the point at which one arrives if one travels from $(1, 0)$ in the clockwise direction for a distance $-\theta$. We begin by showing the sin and cos are differentiable at 0 and that $\sin' 0 = 1$ and $\cos' 0 = 0$. Assume that $0 < \theta < 1$. Then (cf. the figure below), because the shortest path between two point is a line, the line segment L given by $t \in [0, 1] \longmapsto (1-t)(1, 0) + t(\cos\theta, \sin\theta)$ has length less than θ, and therefore, because $t \in [0, 1] \longmapsto (1-t)(\cos\theta, 0) + t(\cos\theta, \sin\theta)$ is one side of a right triangle in which L is the hypotenuse, $\sin\theta \leq \theta$. Since $0 \leq \cos\theta \leq 1$, we now see that $0 \leq 1 - \cos\theta \leq 1 - \cos^2\theta = \sin^2\theta \leq \theta^2$, which, because $\cos(-\theta) = \cos\theta$, proves that cos is differentiable at 0 and $\cos' 0 = 0$. Next consider the region consisting of the right triangle whose vertices are $(0, 0)$, $(\cos\theta, 0)$, and $(\cos\theta, \sin\theta)$ and the rectangle whose vertices are $(\cos\theta, 0)$, $(1, 0)$, $(1, \sin\theta)$, and $(\cos\theta, \sin\theta)$.



derivative of sine

This region contains the wedge cut out of the disk by the horizontal axis and the ray from the origin to $(\cos\theta, \sin\theta)$, and therefore its area is at least as large as that of the wedge. Because the area of the whole disk is π and the arc cut out by the wedge is $\frac{\theta}{2\pi}$th of the whole circle, the area of the wedge is $\frac{\theta}{2}$. Hence, the area of the region is at least $\frac{\theta}{2}$. Since the area of the triangle is $\frac{\sin\theta\cos\theta}{2}$ and that of the rectangle is $(1-\cos\theta)\sin\theta$, this means that

$$\frac{\theta}{2} \le \frac{\sin\theta\cos\theta}{2} + (1-\cos\theta)\sin\theta \le \frac{\sin\theta}{2} + \theta^3$$

and therefore that $1 - 2\theta^2 \le \frac{\sin\theta}{\theta} \le 1$ when $\theta \in (0, 1)$, which, because $\sin(-\theta) = -\sin\theta$, proves that $1 - 2\theta^2 \le \frac{\sin\theta}{\theta} \le 1$ when $\theta \in (-1, 1)$. Thus sin is differentiable at 0 and $\sin' 0 = 1$.

To show that sin and cos are differentiable on $\mathbb{R}$, we will use the trigonometric identities

$$\begin{aligned} \sin(\theta_1 + \theta_2) &= \sin\theta_1\cos\theta_2 + \sin\theta_2\cos\theta_1 \\ \cos(\theta_1 + \theta_2) &= \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2. \end{aligned} \tag{1.5.1}$$

From the first of these we have

$$\frac{\sin(\theta + h) - \sin\theta}{h} = \frac{\sin h}{h}\cos\theta - \sin\theta\frac{1-\cos h}{h},$$

which, by what we already know, tends to $\cos\theta$ as $h \to 0$. Similarly, from the second,

$$\frac{\cos(\theta + h) - \cos\theta}{h} = \cos\theta\frac{\cos h - 1}{h} - \sin\theta\frac{\sin h}{h} \longrightarrow -\sin\theta$$

as $h \to 0$. Thus, sin and cos are differentiable on $\mathbb{R}$ and

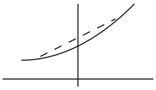
$$\sin'\theta = \cos\theta \quad \text{and} \quad \cos'\theta = -\sin\theta. \tag{1.5.2}$$

1.6 Convex Functions

If I is an interval and $f : I \longrightarrow \mathbb{R}$, then f is said to be *convex* if

$$f\big(\theta x + (1-\theta)y\big) \le \theta f(x) + (1-\theta)f(y) \quad \text{for all } x,\ y \in I \text{ and } \theta \in [0,1].$$

That is, f is convex means that for any pair of points on its graph the line segment joining those points lies above the portion of its graph between them as in:



graph of convex function

It is useful to observe that if $N \ge 2$ and $x_1, \dots, x_N \in I$, then $\sum_{m=1}^N \theta_m x_m \in I$ for all $\theta_1, \dots, \theta_N \in [0,1]$ with $\sum_{m=1}^N \theta_m = 1$. Indeed, this is trivial when $N = 2$. Now assume it is true for N, and let $x_1, \dots, x_{N+1} \in I$ and $\theta_1, \dots, \theta_{N+1} \in [0,1]$ be given. If $\theta_{N+1} = 1$, and therefore $\theta_m = 0$ for $1 \le m \le N$, there is nothing to do. If $\theta_{N+1} < 1$, set $\eta_m = \frac{\theta_m}{1-\theta_{N+1}}$ for $1 \le m \le N$. Then, by assumption, $y \equiv \sum_{m=1}^N \eta_m x_m \in I$, and so

$$\sum_{m=1}^{N+1} \theta_m x_m = (1-\theta_{N+1})y + \theta_{N+1}x_{N+1} \in I.$$

Hence, by induction, we now know the result for all $N \ge 2$. Essentially the same induction procedure shows that if f is convex on I, then

$$f\left(\sum_{m=1}^N \theta_m x_m\right) \le \sum_{m=1}^N \theta_m f(x_m).$$

Lemma 1.6.1 *If f is a continuous function on I, then it is convex if and only if*

$$f\big(2^{-1}x + 2^{-1}y\big) \le 2^{-1}f(x) + 2^{-1}f(y) \quad \textit{for all } x,\ y \in I.$$

Proof We will use induction to prove that, for any $n \ge 1$ and $x_1, \dots, x_{2^n} \in I$,

$$(*) \qquad f\left(\sum_{m=1}^{2^n} 2^{-n}x_m\right) \le 2^{-n}\sum_{m=1}^{2^n} f(x_m).$$

There is nothing to do when $n = 1$. Thus assume that $(*)$ holds for some $n \ge 1$, and let $x_1, \dots, x_{2^{n+1}} \in I$ be given. Note that

$$\sum_{m=1}^{2^{n+1}} 2^{-n-1} x_m = 2^{-1} y + 2^{-1} z$$

$$\text{where } y = \sum_{m=1}^{2^n} 2^{-n} x_m \in I \text{ and } z = \sum_{m=1}^{2^n} 2^{-n} x_{m+2^n} \in I.$$

Hence, $f\left(\sum_{m=1}^{2^{n+1}} 2^{-n-1} x_m\right) \le 2^{-1} f(y) + 2^{-1} f(z)$. Now apply the induction hypothesis to y and z.

Knowing $(*)$, we have that $f\big(\theta x + (1-\theta)y\big) \le \theta f(x) + (1-\theta) f(y)$ for all $x, y \in I$ and all θ of the form $k2^{-n}$ for $n \ge 1$ and $0 \le k \le 2^n$. To see this, take $x_m = x$ if $1 \le m \le k$ and $x_m = y$ if $k+1 \le m \le 2^n$. Then, by $(*)$, one has that $f\big(k2^{-n}x + (1-k2^{-n})y\big) \le k2^{-n} f(x) + (1-k2^{-n}) f(y)$.

Finally, because both $\theta \rightsquigarrow f\big(\theta x + (1-\theta)\big)$ and $\theta \rightsquigarrow \theta f(x) + (1-\theta) f(y)$ are continuous functions on $[0, 1]$, the inequality when θ has the form $k2^{-n}$ implies the inequality for all $\theta \in [0, 1]$. Indeed, given $\theta \in (0, 1)$ and $n \ge 0$, set

$$k = \max\{j \in \mathbb{N} : j \le 2^n \theta\},$$

and observe that $0 \le \theta - k2^{-n} < 2^{-n}$. □

Lemma 1.6.2 *Assume that f is a convex function on the open interval I. If $x \in I$, then*[2]

$$D^+ f(x) \equiv \lim_{y \searrow x} \frac{f(y) - f(x)}{y - x} \text{ and } D^- f(x) \equiv \lim_{y \nearrow x} \frac{f(y) - f(x)}{y - x} \text{ exist in } \mathbb{R}.$$

Moreover, $D^- f(x) \le D^+ f(x)$, and, if equality holds, then f is differentiable at x and $f'(x) = D^{\pm} f(x)$. Finally, if $x, y \in I$ and $x < y$, then $D^+ f(x) \le D^- f(y)$. In particular, f is continuous, and both $D^- f$ and $D^+ f$ are non-decreasing functions on I.

Proof All these assertions come from the inequality

$$f(y) \le \frac{z-y}{z-x} f(x) + \frac{y-x}{z-x} f(z) \quad \text{for } x, y, z \in I \text{ with } x < y < z. \tag{1.6.1}$$

To prove (1.6.1), set $\theta = \frac{z-y}{z-x}$, and observe that $y = \theta x + (1-\theta) z$. Thus $f(y) \le \theta f(x) + (1-\theta) f(z)$, which is equivalent to (1.6.1). By subtracting $f(x)$ from both sides of (1.6.1), we get

$$\frac{f(y) - f(x)}{y - x} \le \frac{f(z) - f(x)}{z - x},$$

[2] We will use $y \searrow x$ when y decreases to x. Similarly, $y \nearrow x$ means that y increases to x.

which shows that the function $y \rightsquigarrow \frac{f(y)-f(x)}{y-x}$ is a non-increasing on the set $\{y \in I : y > x\}$. By subtracting $f(y)$ from both sides of (1.6.1), we get

$$\frac{f(y)-f(x)}{y-x} \le \frac{f(z)-f(y)}{z-y} \quad \text{for } x,\ y,\ z \in I \text{ with } x < y < z.$$

Hence, if $a,\ y \in I$ and $a < x < y$, then

$$\frac{f(x)-f(a)}{x-a} \le \frac{f(y)-f(x)}{y-x},$$

which means that $y \rightsquigarrow \frac{f(y)-f(x)}{y-x}$ is bounded below as well as non-increasing on $\{y \in I : y > x\}$. Thus $D^+ f(x)$ exists in $\mathbb{R}$, and essentially the same argument shows that $D^- f(x)$ does also. Further, if $y' < x < y$, then, again by (1.6.1), one sees that $\frac{f(y')-f(x)}{y'-x} \le \frac{f(y)-f(x)}{y-x}$ and therefore that $D^- f(x) \le D^+ f(x)$. In the case when $D^\pm f(x) = a$, for each $\epsilon > 0$ a $\delta > 0$ can be chosen so that

$$\left| \frac{f(y)-f(x)}{y-x} - a \right| < \epsilon \quad \text{for } y \in I \text{ with } 0 < |y-x| < \delta,$$

and so f is differentiable at x and $f'(x) = a$.

Finally, suppose that $x_1,\ y_1,\ y_2,\ x_2 \in I$ and that $x_1 < y_1 < y_2 < x_2$. Then, by (1.6.1),

$$\frac{f(y_1)-f(x_1)}{y_1-x_1} \le \frac{f(y_2)-f(y_1)}{y_2-y_1} \le \frac{f(x_2)-f(y_2)}{x_2-y_2},$$

and so

$$\frac{f(y_1)-f(x_1)}{y_1-x_1} \le \frac{f(y_2)-f(x_2)}{y_2-x_2}.$$

After letting $y_1 \searrow x_1$ and $y_2 \nearrow x_2$, we see that $D^+ f(x_1) \le D^- f(x_2)$. □

Typical examples of convex functions are $f(x) = x^2$ and $f(x) = |x|$ for $x \in \mathbb{R}$. (See Exercise 1.16 below for a criterion with which to test for convexity.) Since both of these are continuous, Lemma 1.6.1 says that it suffices to check that $\left(\frac{x+y}{2}\right)^2 \le \frac{x^2}{2} + \frac{y^2}{2}$ and $\frac{|x+y|}{2} \le \frac{|x|}{2} + \frac{|y|}{2}$. To see the first of these, observe that, because $(x \pm y)^2 \ge 0$, $2|x||y| \le x^2 + y^2$ and therefore that

$$\left(\frac{x+y}{2}\right)^2 = \frac{x^2+2xy+y^2}{4} \le \frac{x^2+y^2}{2}.$$

As for the second, it is an immediate consequence of the triangle inequality $|x+y| \le |x|+|y|$. When $f(x) = x^2$, both $D^+ f(x)$ and $D^- f(x)$ equal $2x$. When $f(x) = |x|$, $D^+ f(x)$ and $D^- f(x)$ are equal to 1 or -1 depending on whether $x > 0$ or $x < 0$. However, $D^+ f(0) = 1$ but $D^- f(0) = -1$.

1.7 The Exponential and Logarithm Functions

Using Lemma 1.4.2, one can justify the extension of the familiar operations on the set $\mathbb{Q}$ of rational numbers to the set $\mathbb{R}$ of real numbers. For example, although we did not discuss it, we already made tacit use of such an extension in order to extend arithmetic operations from $\mathbb{Q}$ to $\mathbb{R}$. Namely, if $x \in \mathbb{Q}$ and $f_x : \mathbb{Q} \longrightarrow \mathbb{Q}$ is given by $f_x(y) = x + y$, then it is clear that $|f_x(y_2) - f_x(y_1)| = |y_2 - y_1|$ and therefore that f_x is uniformly continuous on $\mathbb{Q}$. Hence, there is a unique extension $\bar{f}_x$ of f_x to $\mathbb{R}$ as a continuous function. Furthermore, if $x_1, x_2 \in \mathbb{Q}$ and $y \in \mathbb{R}$, then $|\bar{f}_{x_2}(y) - \bar{f}_{x_1}(y)| = |x_2 - x_1|$, and so, for each $y \in \mathbb{R}$, there is a unique extension $x \in \mathbb{R} \longmapsto F(x, y) \in \mathbb{R}$ of $x \in \mathbb{Q} \longmapsto \bar{f}_x(y) \in \mathbb{R}$ as a continuous function. That is, $F(x, y)$ is continuous with respect to each of its variables and is equal to $x + y$ when $x, y \in \mathbb{Q}$. From these facts, it is easy to check that $(x, y) \in \mathbb{R}^2 \longmapsto F(x, y) \in \mathbb{R}$ has all the properties that $(x, y) \in \mathbb{Q}^2 \longmapsto x + y$ has. In particular, $F(y, x) = F(x, y)$, $F(x, -x) = 0 \iff y = -x$, and $F\big(F(x, y), z\big) = F\big(x, F(y, z)\big)$. For these reasons, one continues to use the notation $F(x, y) = x + y$ for $x, y \in \mathbb{R}$. A similar extension of the multiplication and division operations can be made, and, together, all these extensions satisfy the same conditions as their antecedents.

A more challenging problem is that of extending exponentiation. That is, given $b \in (0, \infty)$, what is the meaning of b^x for $x \in \mathbb{R}$? When $x \in \mathbb{Z}$, the meaning is clear. Moreover, if $q \in \mathbb{Z}_0 \equiv \mathbb{Z} \setminus \{0\}$, then we would like to take $b^{\frac{1}{q}}$ to be the unique element $a \in (0, \infty)$ such that $a^q = b$. To see that such an element exists and is unique, suppose that $q \ge 1$ and observe that $x \in (0, \infty) \longmapsto x^q \in (0, \infty)$ is strictly increasing, continuous, and that $x^q \ge x$ if $x \ge 1$ and $x^q \le x$ if $x \le 1$. Hence, by Theorem 1.3.6, $(0, \infty) = \{x^q : x \in (0, \infty)\}$, and so, by Theorem 1.4.3, we know not only that $b^{\frac{1}{q}}$ exists and is unique, we also know that $b \rightsquigarrow b^{\frac{1}{q}}$ is continuous. Similarly, if $q \le -1$, there exists a unique, continuous choice of $b \rightsquigarrow b^{\frac{1}{q}}$. Notice that, because $b^{mn} = (b^m)^n$ for $m, n \in \mathbb{Z}$,

$$\big((b^{\frac{1}{q_1}})^{\frac{1}{q_2}}\big)^{q_1 q_2} = \Big(\big((b^{\frac{1}{q_1}})^{\frac{1}{q_2}}\big)^{q_2}\Big)^{q_1} = b,$$

and therefore $b^{\frac{1}{q_1 q_2}} = (b^{\frac{1}{q_1}})^{\frac{1}{q_2}}$. Similarly,

$$\big((b^{\frac{1}{q}})^p\big)^q = (b^{\frac{1}{q}})^{pq} = \big((b^{\frac{1}{q}})^q\big)^p = b^p = \big((b^p)^{\frac{1}{q}}\big)^q,$$

and therefore $(b^{\frac{1}{q}})^p = (b^p)^{\frac{1}{q}}$. Now define $b^r = (b^{\frac{1}{q}})^p = (b^p)^{\frac{1}{q}}$ when $r = \frac{p}{q}$, where $p \in \mathbb{Z}$ and $q \in \mathbb{Z}_0$. To see that this definition is a good one (i.e., doesn't depend on the choice of p and q as long as $r = \frac{p}{q}$), we must show that if $n \in \mathbb{Z}_0$, then $(b^{\frac{1}{nq}})^{np} = (b^{\frac{1}{q}})^p$. But

$$(b^{\frac{1}{nq}})^{np} = \Big(\big((b^{\frac{1}{q}})^{\frac{1}{n}}\big)^n\Big)^p = (b^{\frac{1}{q}})^p.$$

Now that we know how to define b^r for rational r, we need to check that it satisfies the relations

$$(*) \qquad b^{r_1 r_2} = (b^{r_1})^{r_2}, \quad (b_1 b_2)^r = b_1^r b_2^r, \quad \text{and} \quad b^{r_1+r_2} = b^{r_1} b^{r_2}.$$

To check the first of these, observe that

$$(b^{\frac{p_1}{q_1}})^{\frac{p_2}{q_2}} = \Big(\big((b^{p_1})^{\frac{1}{q_1}}\big)^{\frac{1}{q_2}}\Big)^{p_2} = (b^{p_1})^{\frac{p_2}{q_1 q_2}} = \big((b^{p_1})^{p_2}\big)^{\frac{1}{q_1 q_2}} = (b^{p_1 p_2})^{\frac{1}{q_1 q_2}} = b^{\frac{p_1 p_2}{q_1 q_2}}.$$

To prove the second, first note that $\big(b_1^{\frac{1}{q}} b_2^{\frac{1}{q}}\big)^q = b_1 b_2$, and therefore $(b_1 b_2)^{\frac{1}{q}} = b_1^{\frac{1}{q}} b_2^{\frac{1}{q}}$. Hence,

$$(b_1^{\frac{p}{q}})(b_2^{\frac{p}{q}}) = (b_1^{\frac{1}{q}} b_2^{\frac{1}{q}})^p = \big((b_1 b_2)^{\frac{1}{q}}\big)^p = (b_1 b_2)^{\frac{p}{q}}.$$

Finally,

$$(b^{\frac{p_1}{q_1}} b^{\frac{p_2}{q_2}})^{\frac{q_1 q_2}{p_1 q_2 + p_2 q_1}} = (b^{p_1 q_2} b^{p_2 q_1})^{\frac{1}{p_1 q_2 + p_2 q_1}} = (b^{p_1 q_2 + p_2 q_1})^{\frac{1}{p_1 q_2 + p_2 q_1}} = b,$$

and so $b^{\frac{p_1}{q_1}} b^{\frac{p_2}{q_2}} = b^{\frac{p_1 q_2 + p_2 q_1}{q_1 q_2}} = b^{\frac{p_1}{q_1} + \frac{p_2}{q_2}}$.

We are now in a position to show that there exists a continuous extension $\exp_b$ to $\mathbb{R}$ of $r \in \mathbb{Q} \longmapsto b^r \in \mathbb{R}$. For this purpose, first note that $1^r = 1$ for all $r \in \mathbb{Q}$, and therefore that $\exp_1$ is the constant function 1. Now assume that $b > 1$. Then, by the third part of $(*)$, $b^s - b^r = b^r(b^{s-r} - 1)$ for $r < s$. Hence, if $N \geq 1$ and $s, r \in [-N, N] \cap \mathbb{Q}$, then $|b^s - b^r| \leq b^N |b^{s-r} - 1|$, and so we will know that $r \rightsquigarrow b^r$ is uniformly continuous on $[-N, N] \cap \mathbb{Q}$ once we show that for each $\epsilon > 0$ there exists a $\delta > 0$ such that $|b^r - 1| < \epsilon$ whenever $r \in \mathbb{Q} \cap (-\delta, \delta)$. Furthermore, since $b^r > 1$ and $b^r |b^{-r} - 1| = |1 - b^r|$ and therefore $|b^{-r} - 1| \leq |b^r - 1|$ when $r > 0$, we need only check that $\Delta_r \equiv b^r - 1 < \epsilon$ for sufficiently small $r > 0$. To this end, suppose that $0 < r < \frac{1}{n}$. Then $b^{\frac{1}{n}} = b^{\frac{1}{n} - r} b^r \geq b^r$, and so

$$b \geq (b^r)^n = (1 + \Delta_r)^n = \sum_{m=0}^{n} \binom{n}{m} \Delta_r^m \geq 1 + n\Delta_r,$$

which means that $\Delta_r \leq \frac{b-1}{n}$ if $0 < r < \frac{1}{n}$. Hence, if we choose n so that $\frac{b-1}{n} < \epsilon$, then we can take $\delta = \frac{1}{n}$.

We now know that, for each $N \geq 1$, $r \in [-N, N] \cap \mathbb{Q} \longmapsto b^r \in \mathbb{R}$ has a unique extension as a continuous function on $[-N, N]$. Because on $[-N, N]$ the extension corresponding to $N + 1$ must equal the one corresponding to N, this means that there is a unique continuous extension $\exp_b$ to the whole of $\mathbb{R}$. The assumption that $b > 1$ causes no problem, since we can take $\exp_b(x) = \exp_{\frac{1}{b}}(-x)$ if $b < 1$. Notice that, by continuity, each of the relations in $(*)$ extends. That is

$$exp_b(x_1x_2) = \exp_{\exp_b(x_1)}(x_2), \quad \exp_{b_1b_2}(x) = \exp_{b_1}(x)\exp_{b_2}(x),$$
$$\text{and}\ \exp_b(x_1 + x_2) = \exp_b(x_1)\exp_b(x_2).$$

The *exponential functions* $\exp_b$ are among the most interesting and useful ones in mathematics, and, because they are extensions of and share all the algebraic properties of b^r, one often uses the notation b^x instead of $\exp_b(x)$. We will now examine a few of their analytic properties. The first observation is trivial, namely: depending on whether if $b > 1$ or $b < 1$, $x \rightsquigarrow \exp_b(x)$ is strictly increasing or decreasing and tends to infinity or 0 as $x \to \infty$ and tends to 0 or infinity as $x \to -\infty$. We next note that $\exp_b$ is convex. Indeed,

$$\exp_b(x) + \exp_b(y) - 2\exp_b\big(\tfrac{x+y}{2}\big) = \big(\exp_b(2^{-1}x) - \exp_b(2^{-1}y)\big)^2 \geq 0,$$

and so $\exp_b\big(\frac{x+y}{2}\big) \leq \frac{\exp_b(x)+\exp_b(y)}{2}$. Hence convexity follows immediately from Lemma 1.6.1. By Lemma 1.6.2, we now know that $D^+\exp_b(x)$ and $D^-\exp_b(x)$ exist for all $x \in \mathbb{R}$. At the same time, since $\exp_b(y)-\exp_b(x) = \exp_b(x)\big(\exp_b(y-x)-1\big)$, $D^\pm \exp_b(x) = \big(D^\pm \exp_b(0)\big)\exp_b(x)$. Finally,

$$\frac{\exp_b(-y) - 1}{-y} = \exp_b(-y)\frac{\exp_b(y) - 1}{y} \longrightarrow D^+\exp_b(0) \quad \text{as } y \searrow 0,$$

and therefore $D^\pm \exp_b(x) = \big(D^+\exp_b(0)\big)\exp_b(x)$ for all $x \in \mathbb{R}$, which means that $\exp_b$ is differentiable at all $x \in \mathbb{R}$ and that

$$\exp_b' x = (\log b)\exp_b(x) \quad \text{where } \log b \equiv \frac{d\exp_b}{dx}(0). \tag{1.7.1}$$

The choice of the notation $\log b$ is justified by the fact that, like any logarithm function, $\log(b_1b_2) = \log b_1 + \log b_2$. Indeed,

$$\frac{d\exp_{b_1b_2}}{dx}(0) = \frac{d(\exp_{b_1}\exp_{b_2})}{dx}(0) = \frac{d\exp_{b_1}}{dx}(0) + \frac{d\exp_{b_2}}{dx}(0).$$

Of course, as yet we do not know whether it is the trivial logarithm function, the one that is identically 0. To see that it is non-trivial, suppose that $\log 2 = 0$. Then, for any $\epsilon > 0$, $2^{\frac{1}{n}} - 1 \leq \frac{\epsilon}{n}$ for sufficiently large n. But this would mean that

$$2 \leq \big(1 + \tfrac{\epsilon}{n}\big)^n = 1 + \sum_{m=1}^{n}\binom{n}{m}\frac{\epsilon^m}{n^m} \leq 1 + \sum_{m=1}^{n}\frac{\epsilon^m}{m!} \leq 1 + \epsilon\sum_{m=1}^{\infty}\frac{1}{m!} \quad \text{for } 0 < \epsilon \leq 1,$$

and, since, by the ratio test, $\sum_{m=1}^\infty \frac{1}{m!} < \infty$, this would lead to the contradiction that $2 \leq 1$. Hence, we now know that $\log 2 > 0$. Next note that because $\exp_{\exp_b(\lambda)}(x) = \exp_b(\lambda x)$,

$$\log\big(\exp_b(\lambda)\big) = \lambda \log b. \tag{1.7.2}$$

In particular, $\log\bigl(\exp_2(\lambda)\bigr) = \lambda \log 2$, and so if

$$e \equiv \exp_2\left(\frac{1}{\log 2}\right),$$

then $\log e = 1$. If we now define $\exp = \exp_e$, then we see from (1.7.2) and (1.7.1) that

$$\exp'(x) = \exp(x) \quad \text{for all } x \in \mathbb{R}. \tag{1.7.3}$$

In addition, $\log\bigl(\exp(x)\bigr) = x \log e = x$. That is, log is the inverse of exp and, by Theorem 1.4.3, this means that log is continuous. Notice that, because $\exp_{\exp y}(x) = \exp(xy)$ and $b = \exp(\log b)$,

$$b^x = \exp_b(x) = \exp(x \log b) \quad \text{for all } b \in (0,\infty) \text{ and } x \in \mathbb{R}. \tag{1.7.4}$$

Euler made systematic use of the number e, and that accounts for the choice of the letter "e" to denote it. It should be observed that even though we represented e as $\exp_2 \frac{1}{\log 2}$, (1.7.4) shows that

$$e = \exp_b\left(\frac{1}{\log b}\right) \quad \text{for any } b \in (0,\infty) \setminus \{1\}.$$

Before giving another representation, we need to know that

$$\lim_{x \to 0} \frac{\log(1+x)}{x} = 1. \tag{1.7.5}$$

To prove this, note that since log is continuous and vanishes only at 1, and because the derivative of exp is 1 at 0,

$$\frac{x}{\log(1+x)} = \frac{\exp\bigl(\log(1+x)\bigr) - 1}{\log(1+x)} \longrightarrow 1 \text{ as } x \to 0.$$

As an application, it follows that

$$n \log\bigl(1 + \tfrac{x}{n}\bigr) = x \frac{\log\bigl(1 + \frac{x}{n}\bigr)}{\frac{x}{n}} \longrightarrow x \text{ as } n \to \infty,$$

and therefore

$$\bigl(1 + \tfrac{x}{n}\bigr)^n = \exp\Bigl(n \log\bigl(1 + \tfrac{x}{n}\bigr)\Bigr) \longrightarrow \exp(x) \text{ as } n \to \infty.$$

Hence we have shown that

$$\lim_{n\to\infty}\left(1+\tfrac{x}{n}\right)^n = \exp(x) = e^x \quad \text{for all } x \in \mathbb{R}. \tag{1.7.6}$$

In particular, $e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$. Although we now have two very different looking representations of e, neither of them is very helpful when it comes to estimating e. We will get an estimate based on (1.8.4) in the following section.

The preceding computation also allows us to show that log is differentiable and that

$$\log' x = \frac{1}{x} \quad \text{for } x \in (0,\infty). \tag{1.7.7}$$

To see this, note that $\log y - \log x = \log\frac{y}{x} = \log\left(1 + (\frac{y}{x} - 1)\right)$, and therefore, by (1.7.5), that

$$\frac{\log y - \log x}{y - x} = \frac{1}{x}\,\frac{\log\left(1 + (\frac{y}{x} - 1)\right)}{\frac{y}{x} - 1} \longrightarrow \frac{1}{x} \quad \text{as } y \to x.$$

See Exercise 3.8 for another approach, based on (1.7.7), to the logarithm function.

1.8 Some Properties of Differentiable Functions

Suppose that f is a differentiable function on an open set G. If f achieves its maximum value at the point $x \in G$, then $f'(x) = 0$. Indeed,

$$\pm f'(x) = \pm\lim_{h\searrow 0}\frac{f(x\pm h) - f(x)}{\pm h} = \lim_{h\searrow 0}\frac{f(x\pm h) - f(x)}{h} \le 0,$$

and therefore $f'(x) = 0$. Obviously, if f achieves its minimum value at $x \in G$, then the preceding applied to $-f$ shows that again $f'(x) = 0$. This simple observation has many practical consequences. For example, it shows that in order to find the points at which f is at its maximum or minimum, one can restrict one's attention to points at which f' vanishes, a fact that is often called the *first derivative test*.

A more theoretical application is the following theorem.

Theorem 1.8.1 (Mean Value Theorem) *Let f and g be continuous functions on $[a,b]$, and assume that $g(b) \ne g(a)$ and that both functions are differentiable on (a,b). Then there is a $\theta \in (a,b)$ such that*

$$f'(\theta) = g'(\theta)\frac{f(b) - f(a)}{g(b) - g(a)},$$

and so, if $g'(\theta) \ne 0$, then

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(\theta)}{g'(\theta)}.$$

In particular,

$$f(b)-f(a)=f'(\theta)(b-a) \quad \textit{for some } \theta\in(a,b). \tag{1.8.1}$$

Proof Set

$$F(x)=f(x)-\frac{g(b)-g(x)}{g(b)-g(a)}f(a)-\frac{g(x)-g(a)}{g(b)-g(a)}f(b).$$

Then F is differentiable and vanishes at both a and b. Either F is identically 0, in which case F' is also identically 0 and therefore vanishes at every $\theta\in(a,b)$, or there is a point $\theta\in(a,b)$ at which F achieves either its maximum or minimum value, in which case $F'(\theta)=0$. Hence, there always exists some $\theta\in(a,b)$ at which

$$0=F'(\theta)=f'(\theta)+\frac{g'(\theta)f(a)}{g(b)-g(a)}-\frac{g'(\theta)f(b)}{g(b)-g(a)}=f'(\theta)-g'(\theta)\frac{f(b)-f(a)}{g(b)-g(a)},$$

from which the first, and therefore the second, assertion follows. Finally, (1.8.1) follows from the preceding when one takes $g(x)=x$. □

The last part of Theorem 1.8.1 has a nice geometric interpretation in terms of the graph of f. Namely, it says that there is a $\theta\in(a,b)$ at which the slope of the tangent line to the graph at θ is the same as that of the line segment connecting the points $\big(a,f(a)\big)$ and $\big(b,f(b)\big)$.

There are many applications of Theorem 1.8.1. For one thing it says that if $f(b)\neq f(a)$ then there must be a $\theta\in(0,1)$ at which $f'(\theta)\neq 0$. Thus, if $f'=0$ on (a,b), then f is constant on $[a,b]$. Similarly, $f'\geq 0$ on (a,b) if and only if f is non-decreasing, and f must be strictly increasing if $f'>0$ on (a,b). As the function $f(x)=x^3$ shows, the converse of the second of these is not true, since this f is strictly increasing, but its derivative vanishes at 0.

Another application is to a result, known as *L'Hôpital's rule*, which says that if f and g are continuously differentiable, $\mathbb{R}$-valued functions on (a,b) which vanish at some point $c\in(a,b)$ and satisfy $g(x)g'(x)\neq 0$ at $x\neq c$, then

$$\lim_{x\to c}\frac{f(x)}{g(x)}=\lim_{x\to c}\frac{f'(x)}{g'(x)} \quad \text{if } \lim_{x\to c}\frac{f'(x)}{g'(x)} \text{ exists in } \mathbb{R}. \tag{1.8.2}$$

To prove this, apply Theorem 1.8.1 to see that, if $x\in(a,b)\setminus\{c\}$, then

$$\frac{f(x)}{g(x)}=\frac{f(x)-f(c)}{g(x)-g(c)}=\frac{f'(\theta_x)}{g'(\theta_x)}$$

for some θ_x in the open interval between x and c.

Given a function f on an open set G and $n \ge 0$, we use induction to define what it means for f to be *n-times differentiable* on G. Namely, say that any function f on G is 0 times differentiable there, and use $f^{(0)}(x) = f(x)$ to denote its 0th derivative. Next, for $n \ge 1$, say that f is n-times differentiable if it is $(n-1)$ differentiable and $f^{(n-1)}$ is differentiable on G, in which case $f^{(n)}(x) = \partial^n f(x) \equiv \partial f^{(n-1)}(x)$ is its nth derivative. Using induction, it is easy to check the higher order analogs of (1.8.2). That is, if $m \ge 2$ and f and g are m-times continuously differentiable functions that satisfy $f^{(\ell)}(c) = g^{(\ell)}(c) = 0$ for $0 \le \ell < m$ and $g^{(\ell)}(x) \ne 0$ for $0 \le \ell \le m$ and $x \ne c$, then repeated applications of (1.8.2) lead to

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f^{(m)}(x)}{g^{(m)}(x)} \quad \text{if } \lim_{x\to c} \frac{f^{(m)}(x)}{g^{(m)}(x)} \text{ exists in } \mathbb{R}. \tag{1.8.3}$$

For example, consider the functions $1 - \cos x$ and x^2 on $\left(\frac{\pi}{2}, \frac{\pi}{2}\right)$. Then, by (1.5.2) and (1.8.3) with $m = 2$, one has that

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2} = \lim_{x\to 0} \frac{\cos x}{2} = \frac{1}{2}.$$

The following is a very important extension of (1.8.1). Assuming that f is n-times differentiable in an open set containing c, define

$$T_n^f(x; c) = \sum_{m=0}^{n} \frac{f^{(m)}(c)}{m!}(x-c)^m,$$

Notice that if $n \ge 0$, then

$$(*) \qquad \partial T_{n+1}^f(x; c) = T_n^{f'}(x; c).$$

Theorem 1.8.2 (Taylor's Theorem) *If $n \ge 0$ and f is $(n+1)$ times differentiable on $(c-\delta, c+\delta)$ for some $c \in \mathbb{R}$ and $\delta > 0$, then for each $x \in (c-\delta, c+\delta)$ different from c there exists a θ_x in the open interval between c and x such that*

$$\begin{aligned} f(x) &= \sum_{m=0}^{n} \frac{f^{(m)}(c)}{m!}(x-c)^m + \frac{f^{(n+1)}(\theta_x)}{(n+1)!}(x-c)^{n+1} \\ &= T_n^f(x) + \frac{f^{(n+1)}(\theta_x)}{(n+1)!}(x-c)^{n+1}. \end{aligned}$$

Proof When $n = 0$ this is covered by (1.8.1). Now assume that it is true for some $n \ge 0$, and set $F(x) = f(x) - T_{n+1}^f(x; c)$. By the first part of Theorem 1.8.1 (applied with $f = F$ and $g(x) = (x-c)^{n+2}$) and $(*)$, there is a η_x in the open interval between c and x such that

$$\frac{f(x) - T_{n+1}^f(x; c)}{(x-c)^{n+2}} = \frac{f'(\eta_x) - T_n^{f'}(\eta_x; c)}{(n+2)(\eta_x - c)^{n+1}}.$$

By the assumed result for n applied to f', the numerator on the right hand side equals

$$\frac{f^{(n+2)}(\theta_x)}{(n+1)!}(\eta_x - c)^{n+1} \quad \text{for some } \theta_x \text{ in the open interval between } c \text{ and } \eta_x.$$

Hence, $f(x) - T^f_{n+1}(x; c) = \frac{f^{(n+2)}(\theta_x)}{(n+2)!}(x-c)^{n+2}$. □

The polynomial $x \rightsquigarrow T^f_n(x; c)$ is called the nth order *Taylor polynomial* for f at c, and the difference $f - T^f_n(\,\cdot\,; c)$ is called the nth order *remainder*. Obviously, if f is infinitely differentiable and the remainder term tends to 0 as $n \to \infty$, then

$$f(x) = \lim_{n\to\infty} \sum_{m=0}^{n} \frac{f^{(m)}(c)}{m!}(x-c)^m = \sum_{m=0}^{\infty} \frac{f^{(m)}(c)}{m!}(x-c)^m.$$

To see how powerful a result Taylor's theorem is, first note that the binomial formula is a very special case. Namely, take $f(x) = (1+x)^n$, note that $f^{(m)}(x) = n(n-1)\dots(n-m+1)(1+x)^{n-m} = \frac{n!}{(n-m)!}(1+x)^{n-m}$ for $0 \le m \le n$ and $f^{(n+1)} = 0$ everywhere, and conclude from Taylor's theorem that

$$(a+b)^n = a^n\left(1 + \tfrac{b}{a}\right)^n = \sum_{m=0}^{n} \binom{n}{m} a^m b^{n-m} \quad \text{if } a \neq 0.$$

Next observe that the formula $\sum_{m=0}^{\infty} x^m = \frac{1}{1-x}$ for the sum the geometric series with $x \in (-1, 1)$ comes from Taylor's theorem applied to $f(x) = \frac{1}{1-x}$ on $(-1, 1)$. Indeed, for this f, $f^{(m)}(0) = m!$, and therefore Taylor's theorem says that, for each $n \ge 0$

$$\frac{1}{1-x} = \sum_{m=0}^{n} x^m + \theta_x^{n+1} \quad \text{for some} \theta_x \text{ with } |\theta_x| < |x|.$$

Hence, since $|x| < 1$, we get the geometric formula by letting $n \to \infty$.

A more interesting example is the one when $f = \exp$. As we saw in (1.7.3), $\exp' = \exp$, and from this it follows that exp is infinitely differentiable and that $\exp^{(m)} = \exp$ for all $m \ge 0$. Hence, for each $n \ge 0$,

$$e^x = \sum_{m=0}^{n} \frac{x^m}{m!} + \frac{x^{n+1} e^{\theta_x}}{(n+1)!} \quad \text{for some } |\theta_x| < |x|.$$

Since $e^{\theta_x} \le e^{|x|}$ and

$$\left| \frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}} \right| = \frac{|x|}{n+1} \le \frac{1}{2} \quad \text{when } n \ge 2|x|,$$

the last term on the right tends to 0 as $n \to \infty$, and therefore we have now proved that

$$e^x = \sum_{m=0}^{\infty} \frac{x^m}{m!} \quad \text{for all } x \in \mathbb{R}. \tag{1.8.4}$$

The formula in (1.8.4) can also be derived from (1.7.6) and the binomial formula. To see this, apply the binomial formula to see that

$$\left(1 + \tfrac{x}{n}\right)^n = \sum_{m=1}^{n} \frac{x^m}{m!} \frac{\prod_{\ell=0}^{m-1}(n-\ell)}{n^m} = \sum_{m=0}^{n} \frac{x^m}{m!} \prod_{\ell=0}^{m-1}\left(1 - \tfrac{\ell}{n}\right),$$

and so

$$\left|\sum_{m=0}^{n} \frac{x^m}{m!} - \left(1 + \tfrac{x}{n}\right)^n\right| \le \sum_{m=0}^{n} \frac{|x|^m}{m!}\left(1 - \prod_{\ell=0}^{m-1}\left(1 - \tfrac{\ell}{n}\right)\right).$$

For any $N \ge 1$ and $n \ge N$,

$$\sum_{m=0}^{n} \frac{|x|^m}{m!}\left(1 - \prod_{\ell=0}^{m-1}\left(1 - \tfrac{\ell}{n}\right)\right) \le \sum_{m=0}^{N} \frac{|x|^m}{m!}\left(1 - \prod_{\ell=0}^{m-1}\left(1 - \tfrac{\ell}{n}\right)\right) + \sum_{m=N+1}^{\infty} \frac{|x|^m}{m!},$$

and therefore, for all $N \ge 1$,

$$\left|\sum_{m=0}^{\infty} \frac{x^m}{m!} - e^x\right| = \lim_{n\to\infty}\left|\sum_{m=0}^{n} \frac{x^m}{m!} - \left(1 - \tfrac{x}{n}\right)^n\right| \le \sum_{m=N+1}^{\infty} \frac{|x|^m}{m!}.$$

Hence, since, by the ratio test, $\sum_{m=0}^{\infty} \frac{|x|^n}{m!} < \infty$, we arrive again at (1.8.4) after letting $N \to \infty$.

However one arrives at it, (1.8.4) can be used to estimate e. Indeed, observe that for any $m \ge k \ge 2$, $m! \ge (k-1)!k^{m+1-k}$. Hence, since $\sum_{m=k}^{\infty} k^{k-m-1} = \sum_{m=1}^{\infty} k^{-m} = \frac{1}{k-1}$, we see that

$$\sum_{m=0}^{k-1} \frac{1}{m!} \le e \le \sum_{m=0}^{k-1} \frac{1}{m!} + \frac{1}{(k-1)!(k-1)}.$$

Taking $k = 6$, this yields $\frac{163}{60} \le e \le \frac{1631}{600}$. Since $\frac{163}{60} \ge 2.71$ and $\frac{1631}{600} \le 2.72$, $2.71 \le e \le 2.72$. Taking larger k's, one finds that e is slightly larger than 2.718. Nonetheless, e is very far from being a rational number. In fact, it was the first number that was shown to be transcendental (i.e., no non-zero polynomial with integer coefficients vanishes at it).

We next apply the same sort of analysis to the functions sin and cos. Recall (cf. (1.5.2)) that $\sin' = \cos$ and $\cos' = -\sin$. Hence, $\sin^{(2m)} = (-1)^m \sin$, $\sin^{(2m+1)} = (-1)^m \cos$, $\cos^{(m)} = (-1)^m \cos$, and $\cos^{(2m+1)} = (-1)^{m+1} \sin$.

Since $\sin 0 = 0$ and $\cos 0 = 1$, we see that $\sin^{(2m)} 0 = 0$, $\sin^{(2m+1)} 0 = (-1)^m$, $\cos^{(2m)} 0 = (-1)^m$, and $\cos^{(2m+1)} 0 = 0$. Thus, just as in the preceding,

$$\sin x = \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m+1}}{(2m+1)!} \quad \text{and} \quad \cos x = \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m}}{(2m)!}. \tag{1.8.5}$$

Finally, we turn to log. Remember (cf. (1.7.7)) that $\log' x = \frac{1}{x}$, and therefore that $\log'(1-x) = -\frac{1}{1-x}$ for $|x| < 1$. Using our computation when we derived the geometric formula, we now see that the mth derivative $-\frac{d^m}{dx^m} \log(1-x)$ of $-\log(1-x)$ at $x = 0$ is $(m-1)!$ when $m \ge 1$. Since $\log 1 = 0$, Taylor's theorem says that, for each $n \ge 0$ there exists a $|\theta_x| < |x|$ such that

$$-\log(1-x) = \sum_{m=1}^{n} \frac{x^m}{m} + \frac{\theta_x^{n+1}}{n+1}.$$

Hence, we have that

$$\log(1-x) = -\sum_{m=1}^{\infty} \frac{x^m}{m} \quad \text{for } x \in (-1, 1). \tag{1.8.6}$$

It should be pointed out that there are infinitely differentiable functions for which Taylor's theorem gives essentially no useful information. To produce such an example, we will use the following corollary of Theorem 1.8.1.

Lemma 1.8.3 *Assume that $a < c < b$ and that $f : (a, b) \longrightarrow \mathbb{R}$ is a continuous function that is differentiable on $(a, c) \cup (c, b)$. If $\alpha = \lim_{x\to c} f'(x)$ exists in $\mathbb{R}$, then f is differentiable at c and $f'(c) = \alpha$.*

Proof For $x \in (c, b)$, Theorem 1.8.1 says that $\frac{f(x)-f(c)}{x-c} = f'(\theta_x)$ for some $\theta_x \in (c, x)$. Hence, as $x \searrow c$, $\frac{f(x)-f(c)}{x-c} \longrightarrow \alpha$. Similarly, $\frac{f(x)-f(c)}{x-c} \longrightarrow \alpha$ as $x \nearrow c$, and therefore $\lim_{x\to c} \frac{f(x)-f(c)}{x-c} = \alpha$. □

Now consider the function f given by $f(x) = e^{-\frac{1}{|x|}}$ for $x \neq 0$ and 0 when $x = 0$. Obviously, f is continuous on $\mathbb{R}$ and infinitely differentiable on $(-\infty, 0) \cup (0, \infty)$. Furthermore, by induction one sees that there are $2m$th order polynomials $P_{m,+}$ and $P_{m,-}$ such that

$$f^{(m)}(x) = \begin{cases} P_{m,+}(x^{-1}) f(x) & \text{if } x > 0 \\ P_{m,-}(x^{-1}) f(x) & \text{if } x < 0. \end{cases}$$

Since, by (1.8.4), $e^x \ge \frac{x^{2m+1}}{(2m+1)!}$ for $x \ge 0$, it follows that

$$\lim_{x\searrow 0} f^{(m)}(x) = \lim_{x\nearrow\infty} \frac{P_{m,+}(x)}{e^x} = 0.$$

Similarly, $\lim_{x\nearrow 0} f^{(m)}(x) = 0$. Hence, by Lemma 1.8.3, f has continuous derivatives of all orders and all of them vanish at 0. As a result, when we apply Taylor's theorem to f at 0, all the Taylor polynomials vanish and therefore the remainder term is equal to f. Since the whole point of Taylor's theorem is that the Taylor polynomial should be the dominant term nearby the place where the expansion is made, we see that it gives no useful information at 0 when applied to this function.

A slightly different sort of application of Theorem 1.8.2 is the following. Suppose that f is a continuously differentiable function and that its derivative is tending rapidly to a finite constant as $x \to \infty$. Then $f(n+1) - f(n)$ will be approximately equal to $f'(n)$ for large n's, and so $f(n+1) - f(0)$ will be approximately equal to $\sum_{n=0}^{n} f'(m)$. Applying this idea to the function $f(x) = \log(1+x)$, we are guessing that $\log(n+1)$ is approximately equal to $\sum_{m=1}^{n} \frac{1}{m}$. To check this, note that, by Taylor's theorem, for $n \geq 1$,

$$\log(n+1) - \log n = \log\bigl(1 + \tfrac{1}{n}\bigr) = \frac{1}{n} - \frac{1}{2(1+\theta_n)^2 n^2} \quad \text{for some } 0 < \theta_n < \frac{1}{n}.$$

Therefore, if $\Delta_0 = 0$ and $\Delta_n = \sum_{m=1}^{n} \frac{1}{m} - \log(n+1)$ for $n \geq 1$, then

$$0 < \Delta_n - \Delta_{n-1} \leq \frac{1}{2n^2} \quad \text{for } n \geq 1.$$

Hence $\{\Delta_n : n \geq 1\}$ is a strictly increasing sequence of positive numbers that are bounded above by $C = \frac{1}{2}\sum_{m=1}^{\infty} \frac{1}{m^2}$, and as such they converge to some $\gamma \in (0, C]$. In other words

$$0 < \sum_{m=1}^{n} \frac{1}{m} - \log(n+1) \nearrow \gamma \quad \text{as } n \to \infty.$$

This result was discovered by Euler, and the constant γ is sometimes called *Euler's constant*. What it shows is that *the partial sums of the harmonic series grow logarithmically*.

A more delicate example of the same sort is the one discovered by DeMoivre. What he wanted to know is how fast $n!$ grows. For this purpose, he looked at $\log(n!) = \sum_{m=1}^{n} \log m$. Since $\log x = f'(x)$ when $f(x) = x \log x - x$, the preceding reasoning suggests that $\log(n!)$ should grow approximately like $n \log n - n$. However, because the first derivative of this f is not approaching a finite constant at ∞, one has to look more closely. By Theorem 1.8.1, $f(n+1) - f(n) = \log \theta_n$ for some $\theta_n \in [n, n+1]$. Thus, $f(n+1) - f(n) \geq \log n$ and $f(n) - f(n-1) \leq \log n$, and so, with luck, the average $\frac{f(n+1)+f(n)}{2}$ should be a better approximation of $\log(n!)$ than $f(n)$ itself. Note that

$$\frac{f(n+1)+f(n)}{2} = \bigl(n + \tfrac{1}{2}\bigr)\log n - n + \tfrac{1}{2}\bigl((n+1)\log(1+\tfrac{1}{n}) - 1\bigr)$$

and (cf. (1.7.5)) that the last expression tends to 0 as $n \to \infty$. Hence, we are now led to look at

$$\Delta_n \equiv \log(n!) - \big(n + \tfrac{1}{2}\big)\log n + n.$$

Clearly, $\Delta_{n+1} - \Delta_n$ equals

$$\log(n+1) - \big(n + \tfrac{3}{2}\big)\log(n+1) + \big(n + \tfrac{1}{2}\big)\log n + 1 = 1 - \big(n + \tfrac{1}{2}\big)\log\big(1 + \tfrac{1}{n}\big).$$

By Taylor's theorem, $\log\big(1 + \frac{1}{n}\big) = \frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3(1+\theta_n)^3 n^3}$ for some $\theta_n \in \big(0, \frac{1}{n}\big)$, and therefore

$$\big(n + \tfrac{1}{2}\big)\log\big(1 + \tfrac{1}{n}\big) = 1 - \frac{1}{4n^2} + \frac{2n+1}{6(1+\theta_n)^3 n^3}.$$

Hence, we now know that

$$-\frac{1}{2n^2} \le -\frac{2n+1}{6(1+\theta_n)^3 n^3} \le \Delta_{n+1} - \Delta_n \le \frac{1}{4n^2},$$

and therefore that

$$-\frac{1}{2}\sum_{m=n_1}^{n_2-1}\frac{1}{m^2} \le \Delta_{n_2} - \Delta_{n_1} \le \frac{1}{4}\sum_{m=n_1}^{n_2-1}\frac{1}{m^2} \quad \text{for all } 1 \le n_1 < n_2.$$

Since $\frac{1}{m^2} \le \frac{2}{m(m+1)} = 2\big(\frac{1}{m} - \frac{1}{m+1}\big)$, $\sum_{m=n_1}^{n_2-1}\frac{1}{m^2} \le 2\big(\frac{1}{n_1} - \frac{1}{n_2}\big) \le \frac{2}{n_1}$, which means that

$$|\Delta_{n_2} - \Delta_{n_1}| \le \frac{1}{n_1} \quad \text{for } 1 \le n_1 < n_2.$$

By Cauchy's criterion, it follows that $\{\Delta_n : n \ge 1\}$ converges to some $\Delta \in \mathbb{R}$ and that $|\Delta - \Delta_n| = \lim_{\ell\to\infty}|\Delta_\ell - \Delta_n| \le \frac{1}{n}$. Because $\Delta_n = \log\frac{n!e^n}{n^{n+\frac{1}{2}}}$, this is equivalent to

$$e^{\Delta - \frac{1}{n}} \le \frac{n!e^n}{n^{n+\frac{1}{2}}} \le e^{\Delta + \frac{1}{n}} \quad \text{for all } n \ge 1. \tag{1.8.7}$$

In other words, in that their ratio is caught between $e^{-\frac{2}{n}}$ and $e^{\frac{2}{n}}$, $n!$ is growing very much like $e^{\Delta}\sqrt{n}\,\big(\frac{n}{e}\big)^n$. In particular,

$$\lim_{n\to\infty}\frac{n!}{e^{\Delta}\sqrt{n}\,\big(\frac{n}{e}\big)^n} = 1,$$

Such a limit result is called an *asymptotic limit* and is often abbreviated by $n! \sim e^{\Delta}\sqrt{n}\,\big(\frac{n}{e}\big)$. This was the result proved by DeMoivre. Shortly thereafter, Stirling showed that $e^{\Delta} = \sqrt{2\pi}$ (cf. (3.2.4) and (5.2.5) below), and ever since the result has (somewhat unfairly) been known as *Stirling's formula*.

It is worth thinking about the difference between this example and the preceding one. In both examples we needed $\Delta_{n+1} - \Delta_n$ to be of order n^{-2}. In the first example we did this by looking at the first order Taylor polynomial for $\log(1+x)$ and showing that it gets canceled, leaving us with terms of order n^{-2}. In the second example, we needed to cancel the first two terms before being left with terms of order n^{-2}, and it was in the cancellation of the second term that the use of $(n+\frac{1}{2})\log(1+\frac{1}{n})$ instead of $n\log(1+\frac{1}{n})$ played a critical role.

One often wants to look at functions that are obtained by composing one function with another. That is, if a function f takes values where the function g is defined, then the *composite function* $g \circ f$ is the one for which $g \circ f(x) = g\big(f(x)\big)$ (cf. Exercise 1.9). Thinking of g as a machine that produces an output from its input and of $f(x)$ as measuring the quantity of input at time x, it is reasonable to think that the derivative of $g \circ f$ should be the product of the rate at which the machine produces its output times the rate at which input is being fed into the machine, and the first part of the following theorem shows that this expectation is correct.

Theorem 1.8.4 *Suppose that $f : (a,b) \longrightarrow (c,d)$ and $g : (c,d) \longrightarrow \mathbb{R}$ are continuously differentiable functions. Then $g \circ f$ is continuously differentiable and*

$$(g \circ f)'(x) = (g' \circ f)(x) f'(x) \quad \text{for } x \in (a,b).$$

Next, assume that $f : [a,b] \longrightarrow \mathbb{R}$ is continuous and continuously differentiable on (a,b). If $f' > 0$, then f has an inverse f^{-1}, f^{-1} is continuously differentiable on $\big(f(a), f(b)\big)$, and

$$\frac{df^{-1}}{dy}(y) = \frac{1}{(f' \circ f^{-1})(y)} \quad \text{for all } y \in \big(f(a), f(b)\big).$$

Proof Given unequal $x_1, x_2 \in (a,b)$, the last part of Theorem 1.8.1 implies that there exists a θ_{x_1,x_2} in the open interval between $f(x_1)$ and $f(x_2)$ such that

$$g \circ f(x_2) - g \circ f(x_1) = g'(\theta_{x_1,x_2})\big(f(x_2) - f(x_1)\big).$$

Hence, as $x_2 \to x_1$,

$$\frac{g \circ f(x_2) - g \circ f(x_1)}{x_2 - x_1} = g'(\theta_{x_1,x_2}) \frac{f(x_2) - f(x_1)}{x_2 - x_1} \longrightarrow (g' \circ f)(x_1) f'(x_1).$$

To prove the second assertion, first note that, by (1.8.1), f is strictly increasing and therefore has a continuous inverse. Now let $y_1, y_2 \in \big(f(a), f(b)\big)$ be unequal points, and apply (1.8.1) to see that

$$y_2 - y_1 = f \circ f^{-1}(y_2) - f \circ f^{-1}(y_1) = f'(\theta_{y_1,y_2})\big(f^{-1}(y_2) - f^{-1}(y_1)\big)$$

for some θ_{y_1,y_2} between $f^{-1}(y_1)$ and $f^{-1}(y_2)$. Hence

$$\frac{f^{-1}(y_2) - f^{-1}(y_1)}{y_2 - y_1} = \frac{1}{f'(\theta_{y_1,y_2})} \longrightarrow \frac{1}{f'\big(f^{-1}(y_1)\big)} \quad \text{as } y_2 \to y_1. \qquad \square$$

The first result in Theorem 1.8.4 is called the *chain rule*. Notice that we could have used the second part to prove that log is differentiable and that $\log' x = \frac{1}{x}$. Indeed, because $\exp' = \exp$ and $\log = \exp^{-1}$, $\log' x = \frac{1}{\exp \circ \log(x)} = \frac{1}{x}$. An application of the chain rule is provided by the functions $x \in (0, \infty) \longmapsto x^\alpha \in (0, \infty)$ for any $\alpha \in \mathbb{R}$. Because $x^\alpha = \exp(\alpha \log x)$, $\frac{dx^\alpha}{dx} = \alpha \exp(\alpha \log x)\frac{1}{x} = \alpha x^{\alpha-1}$.

1.9 Infinite Products

Closely related to questions about the convergence of series are the analogous questions about the convergence of products, and, because the exponential and logarithmic functions play a useful role in their answers, it seemed appropriate to postpone this topic until we had those functions at our disposal.

Given a sequence $\{a_m : m \geq 1\} \subseteq \mathbb{R}$, consider the problem of giving a meaning to their product. The procedure is very much like that for sums. One starts by looking at the partial product $\prod_{m=1}^n a_m = a_1 \times \cdots \times a_n$ of the first n factors, and then asks whether the sequence $\left\{\prod_{m=1}^n a_m : n \geq 1\right\}$ of partial products converges in $\mathbb{R}$ as $n \to \infty$. Obviously, one should suspect that the convergence of these partial products should be intimately related to the rate at which the a_m's are tending to 1. For example, suppose that $a_m = 1 + \frac{1}{m^\alpha}$ for some $\alpha > 0$, and set $P_n(\alpha) = \prod_{m=1}^n a_m$. Clearly $0 \leq P_n(\alpha) \leq P_{n+1}(\alpha)$, and so $\{P_n(\alpha) : n \geq 1\}$ converges in $\mathbb{R}$ if and only if $\sup_{n\geq 1} P_n(\alpha) < \infty$. Furthermore, $\alpha \rightsquigarrow P_n(\alpha)$ is non-increasing for each n, and so

$$\sup_{n\geq 1} P_n(\alpha) = \infty \implies \sup_{n\geq 1} P_n(\beta) = \infty \quad \text{if } 0 < \beta < \alpha.$$

Observe that $\log P_n(1) = n + 1 \longrightarrow \infty$, and so $\sup_{n\geq 1} P_n(\alpha) = \infty$ if $\alpha \in (0, 1]$. On the other hand, if $\alpha > 1$, then

$$\log\left(\prod_{m=1}^n a_m\right) = \sum_{m=1}^n \log\left(1 + \tfrac{1}{m^\alpha}\right) = \sum_{m=1}^n \tfrac{1}{m^\alpha} + \sum_{m=1}^n \left(\log\left(1 + \tfrac{1}{m^\alpha}\right) - \tfrac{1}{m^\alpha}\right).$$

Since, by (1.8.6),

$$\left|\log(1+x) - x\right| = \left|\sum_{k=2}^\infty \frac{(-x)^k}{k}\right| \leq \frac{x^2}{2(1-|x|)} \leq x^2 \quad \text{for } |x| \leq \tfrac{1}{2},$$

it follows that that $\sup_{n\geq 1} P_n(\alpha) < \infty$ if $\alpha > 1$.

There is an annoying feature here that didn't arise earlier. Namely, if any one of the a_m's is 0, then regardless of what the other factors are, the limit will exist and be equal to 0. Because we want convergence to reflect properties that do not depend on any finite number of factors, we adopt the following, somewhat convoluted, definition. Given $\{a_m : m \geq 1\} \subseteq \mathbb{R}$, we will now say that the infinite product $\prod_{m=1}^{\infty} a_m$ *converges* if there exists an m_0 such that $a_m \neq 0$ for $m \geq m_0$ and the sequence $\left\{\prod_{m=m_0}^{n} a_m : n \geq m_0\right\}$ converges to a real number other than 0, in which case we take the smallest such m_0 and define

$$\prod_{m=1}^{\infty} a_m = \begin{cases} 0 & \text{if } m_0 > 1 \\ \lim_{n\to\infty} \prod_{m=1}^{n} a_m & \text{if } m_0 = 1. \end{cases}$$

It will be made clear in Lemma 1.9.1, which is the analog of Cauchy's criterion for series, below why we insist that the limit of non-zero factors not be 0 and say that the product of non-zero factors *diverges to* 0 if the limit of its partial products is 0.

Lemma 1.9.1 *Given* $\{a_m : m \geq 1\} \subseteq \mathbb{R}$, $\prod_{m=1}^{\infty} a_m$ *converges if and only if* $a_m = 0$ *for at most finitely many* $m \geq 1$ *and*

$$\lim_{n_1\to\infty} \sup_{n_2>n_1} \left| \prod_{m=n_1+1}^{n_2} a_m - 1 \right| = 0. \tag{1.9.1}$$

Proof First suppose that $\prod_{m=1}^{\infty} a_m$ converges. Then, without loss in generality, we will assume the $a_m \neq 0$ for any $m \geq 1$ and that there exists an $\alpha > 0$ such that $\left|\prod_{m=1}^{n} a_m\right| \geq \alpha$ for all $n \geq 1$. Hence, for $n_2 > n_1$,

$$\left| \prod_{m=1}^{n_2} a_m - \prod_{m=1}^{n_1} a_m \right| \geq \alpha \left| \prod_{m=n_1+1}^{n_2} a_m - 1 \right|,$$

and therefore, by Cauchy's criterion for the sequence $\left\{\prod_{m=1}^{n} a_m : n \geq 1\right\}$, (1.9.1) holds.

In proving the converse, we may and will again assume that $a_m \neq 0$ for any $m \geq 1$. Set $P_n = \prod_{m=1}^{n} a_m$. By (1.9.1), we know that there exists an n_1 such that $\frac{1}{2} \leq \prod_{m=n_1+1}^{n_2} a_m \leq 2$ and therefore

$$|P_{n_2}| = |P_{n_1}| \left| \prod_{m=n_1+1}^{n_2} a_m \right| \begin{cases} \geq \frac{|P_{n_1}|}{2} \\ \leq 2|P_{n_1}| \end{cases} \quad \text{for } n_2 > n_1,$$

from which it follows that there exists an $\alpha \in (0, 1)$ such that $\alpha \leq |P_n| \leq \frac{1}{\alpha}$ for all $n \geq 1$. Hence, since

$$|P_{n_2} - P_{n_1}| = |P_{n_1}| \left| \prod_{m=n_1+1}^{n_2} a_m - 1 \right| \begin{cases} \leq \frac{1}{\alpha} \left| \prod_{m=n_1+1}^{n_2} a_m - 1 \right| \\ \geq \alpha \left| \prod_{m=n_1+1}^{n_2} a_m - 1 \right|, \end{cases}$$

$\{P_n : n \geq 1\}$ satisfies Cauchy's convergence criterion and its limit cannot be 0. □

By taking $n_2 = n_1 + 1$ in (1.9.1), we see that a necessary condition for the convergence of $\prod_{m=1}^{\infty} a_m$ is that $a_m \longrightarrow 1$ as $m \to \infty$. On the other hand, this is not a sufficient condition. For example, we already showed that $\prod_{m=1}^{n} a_m = n+1 \longrightarrow \infty$ when $a_m = 1 + \frac{1}{m}$, and if $a_m = 1 - \frac{1}{m+1}$, then $\prod_{m=1}^{n} a_m = \frac{1}{n+1} \longrightarrow 0$. The situation here is reminiscent of the one for series, and, as the following lemma shows, it is closely related to the one for series.

Lemma 1.9.2 *Suppose that $a_m = 1 + b_m$, where $b_m \in \mathbb{R}$ for all $m \geq 1$. Then $\prod_{m=1}^{\infty} a_m$ converges if $\sum_{m=1}^{\infty} |b_m| < \infty$. Moreover, if $b_m \geq 0$ for all $m \geq 1$, $\sum_{m=1}^{\infty} b_m < \infty$ if $\prod_{m=1}^{\infty} a_m$ converges.*

Proof Begin by observing that

$$(*) \qquad \prod_{m=n_1+1}^{n_2} (1 + b_m) - 1 = \sum_{\emptyset \neq F \subseteq \mathbb{Z}^+ \cap [n_1+1, n_2]} b_F \quad \text{where } b_F = \prod_{m \in F} b_m.$$

If $b_m \geq 0$ for all $m \geq 1$, this shows that $\prod_{m=1}^{n}(1 + b_m) - 1 \geq \sum_{m=1}^{n} b_m$ for all $n \geq 1$ and therefore that $\sum_{m=1}^{\infty} b_m < \infty$ if $\prod_{m=1}^{\infty}(1 + b_m)$ converges. Conversely, still assuming that $b_m \geq 0$ for all $m \geq 1$,

$$0 \leq \prod_{m=n_1+1}^{n_2} (1 + b_m) - 1 \leq \exp\left(\sum_{m=n_1+1}^{n_2} b_m \right) - 1$$

since $1 + b_m \leq e^{b_m}$, and therefore

$$\sum_{m=1}^{\infty} b_m < \infty \implies \lim_{n_1 \to \infty} \sup_{n_2 > n_1} \left| \prod_{m=n_1+1}^{n_2} (1 + b_m) - 1 \right| = 0,$$

which, by Lemma 1.9.1 with $a_m = 1 + b_m$, proves that $\prod_{m=1}^{\infty}(1 + b_m)$ converges.

Finally, suppose that $\sum_{m=1}^{\infty} |b_m| < \infty$. To show that $\prod_{m=1}^{\infty}(1 + b_m)$ converges, first note that, since $b_m \longrightarrow 0$, $1 + b_m = 0$ for at most a finite number of m's. Next, use $(*)$ to see that

$$\left| \prod_{m=n_1+1}^{n_2} (1 + b_m) - 1 \right| \leq \sum_{\emptyset \neq F \subseteq \mathbb{Z}^+ \cap [n_1+1, n_2]} |b_F| = \left| \prod_{m=n_1+1}^{n_2} (1 + |b_m|) - 1 \right|.$$

Hence, since $\prod_{m=1}^{\infty}(1+|b_m|)$ converges, (1.9.1) holds with $a_m = 1 + b_m$, and so, by Lemma 1.9.1, $\prod_{m=1}^{\infty}(1+b_m)$ converges. □

When $\sum_{m=1}^{\infty}|b_m| < \infty$, the product $\prod_{m=1}^{\infty}(1+b_m)$ is said to be *absolutely convergent* .

1.10 Exercises

Exercise 1.1 Show that

$$\lim_{n\to\infty}\left(n^{\frac{1}{2}}(1+n)^{\frac{1}{2}} - n\right) = \frac{1}{2}.$$

Exercise 1.2 Show that for any $\alpha > 0$ and $a \in (-1, 1)$, $\lim_{n\to\infty} n^{\alpha}a^n = 0$.

Exercise 1.3 Given $\{x_n : n \geq 1\} \subseteq \mathbb{R}$, consider the averages $A_n \equiv \frac{\sum_{m=1}^{n} x_m}{n}$ for $n \geq 1$. Show that if $x_n \longrightarrow x$ in $\mathbb{R}$, then $A_n \longrightarrow x$. On the other hand, construct an example for which $\{A_n : n \geq 1\}$ converges in $\mathbb{R}$ but $\{x_n : n \geq 1\}$ does not.

Exercise 1.4 Although we know that $\sum_{m=1}^{\infty}\frac{1}{m^2}$ converges, it is not so easy to find out what it converges to. It turns out that it converges to $\frac{\pi^2}{6}$, but showing this requires work (cf. (3.4.6) below). On the other hand, show that $\sum_{m=1}^{n}\frac{1}{m(m+1)} = 1 - \frac{1}{n+1} \longrightarrow 1$.

Exercise 1.5 There is a profound distinction between absolutely convergent series and those that are conditionally but not absolutely convergent. The distinction is that the sum of an absolutely convergent series is, like that of a finite sum, the same for all orderings of the summands. More precisely, show that if $\sum_{m=0}^{\infty} a_m$ is absolutely convergent and its sum is s, then for each $\epsilon > 0$ there is a finite set $F_\epsilon \subseteq \mathbb{N}$ such that $\left|s - \sum_{m\in S} a_m\right| < \epsilon$ for all $S \subseteq \mathbb{N}$ containing F_ϵ. As a consequence, show that $s = \sum_{m=0}^{\infty} a_m^+ - \sum_{m=0}^{\infty} a_m^-$.

On the other hand, if $\sum_{m=0}^{\infty} a_m$ is conditionally but not absolutely convergent, show that, for *any* $s \in \mathbb{R}$, there is an increasing sequence $\{F_\ell : \ell \geq 1\}$ of finite subsets of $\mathbb{N}$ such that $\mathbb{N} = \bigcup_{\ell=1}^{\infty} F_\ell$ and $\sum_{m\in F_\ell} a_m \longrightarrow s$ as $\ell \to \infty$. To prove this second assertion, begin by showing that $a_m \longrightarrow 0$ and that both $\sum_{m=0}^{\infty} a_m^+$ and $\sum_{m=0}^{\infty} a_m^-$ are infinite. Next, given $s \in [0, \infty)$, take

$$n_+ = \min\left\{n \geq 0 : \sum_{m=0}^{n} a_m^+ > s\right\},\quad n_- = \min\left\{n \geq 0 : \sum_{m=0}^{n_+} a_m^+ - \sum_{m=0}^{n} a_m^- < s\right\},$$

and define

$$F_1 = \{0 \leq m \leq n_+ : a_m \geq 0\} \text{ and } F_2 = F_1 \cup \{0 \leq m \leq n_- : a_m < 0\}.$$

Proceeding by induction, given $F_1, \dots, F_\ell$ for some $\ell \ge 2$, take

$$n_+ = \min\left\{ n \notin F_\ell : \sum_{m \in F_\ell} a_m + \sum_{\substack{1 \le m \le n \\ m \notin F_\ell}} a_m^+ > s \right\}$$
$$\text{and } F_{\ell+1} = F_\ell \cup \{m \notin F_\ell : 1 \le m \le n_+ \ \& \ a_m \ge 0\}$$

if ℓ is even, and

$$n_- = \min\left\{ n \notin F_\ell : \sum_{m \in F_\ell} a_m - \sum_{\substack{1 \le m \le n \\ m \notin F_\ell}} a_m^- < s \right\}$$
$$\text{and } F_{\ell+1} = F_\ell \cup \{m \notin F_\ell : 1 \le m \le n_- \ \& \ a_m < 0\}$$

if ℓ is odd. Show that $\mathbb{N} = \bigcup_{\ell=1}^{\infty} F_\ell$ and that, for each $\ell \ge 1$, $\left| s - \sum_{m \in F_\ell} a_m \right| \le |a_{n_\ell}|$ where $n_\ell = \max\{n : n \in F_\ell\}$. If $s < 0$, the same argument works, only then one has to get below s on odd steps and above s on even ones. Finally, notice that this line of reasoning shows that if $\{a_m : m \ge 0\} \subseteq \mathbb{R}$ is any sequence having the properties that $a_m \longrightarrow 0$ and $\sum_{m=0}^{\infty} a_m^+ = \sum_{m=0}^{\infty} a_m^- = \infty$, then for each $s \in R$ there is a permutation π of $\mathbb{N}$ such that $\sum_{m=0}^{\infty} a_{\pi(m)}$ converges to s.

Exercise 1.6 Let $\{a_m : m \ge 0\}$ and $\{b_m : m \ge 0\}$ be a pair of sequences in $\mathbb{R}$, and set $S_n = \sum_{m=0}^{n} a_m$. Show that

$$\sum_{m=0}^{n} a_m b_m = S_n b_n + \sum_{m=0}^{n-1} S_m (b_m - b_{m+1}),$$

a formula that is known as the *summation by parts* formula. Next, assume that $\sum_{m=0}^{\infty} a_m$ converges to $S \in \mathbb{R}$, and show that

$$\sum_{m=0}^{\infty} a_m \lambda^m - S = (1 - \lambda) \sum_{m=0}^{\infty} (S_m - S) \lambda^m \quad \text{if } |\lambda| < 1.$$

Given $\epsilon > 0$, choose $n_\epsilon \ge 1$ so that $|S_n - S| < \epsilon$ for $n \ge n_\epsilon$, and conclude first that

$$\left| \sum_{m=0}^{\infty} a_m \lambda^m - S \right| \le (1 - \lambda) \sum_{m=0}^{n_\epsilon - 1} |S_m - S| \lambda^m + \epsilon \quad \text{if } 0 < \lambda < 1$$

and then that

$$\lim_{\lambda \nearrow 1} \sum_{m=0}^{\infty} a_m \lambda^m = \sum_{m=0}^{\infty} a_m. \tag{1.10.1}$$

This observation, which is usually attributed to Abel, requires that one know ahead of time that $\sum_{m=0}^{\infty} a_m$ converges. Indeed, give an example of a sequence for which the limit on the left exists in $\mathbb{R}$ but the series on the right diverges. To see how (1.10.1) can be useful, use it to show that $\sum_{m=1}^{\infty} \frac{(-1)^m}{m} = \log \frac{1}{2}$. A generalization of this idea is given in part (**v**) of Exercise 3.16 below.

Exercise 1.7 Whatever procedure one uses to construct the real numbers starting from the rational numbers, one can represent positive real numbers in terms of D-bit expansions. That is, let $D \geq 2$ be an integer, define Ω be the set of all maps $\omega : \mathbb{N} \longrightarrow \{0, 1, \dots, D-1\}$, and let $\tilde{\Omega}$ denote the set of $\omega \in \Omega$ such that $\omega(0) \neq 0$ and $\omega(k) < D-1$ for infinitely many k's.

(i) Show for any $n \in \mathbb{Z}$ and $\omega \in \tilde{\Omega}$, $\sum_{k=0}^{\infty} \omega(k) D^{n-k}$ converges to an element of $[D^n, D^{n+1})$. In addition, show that for each $x \in [D^n, D^{n+1})$ there is a unique $\omega \in \tilde{\Omega}$ such that $x = \sum_{k=0}^{\infty} \omega(k) D^{n-k}$. Conclude that $\mathbb{R}$ has the same cardinality as $\tilde{\Omega}$.

(ii) Show that $\Omega \setminus \tilde{\Omega}$ is countable and therefore that $\tilde{\Omega}$ is uncountable if and only if Ω is. Next show that Ω is uncountable by the following famous *anti-diagonal* argument devised by Cantor. Suppose that Ω were countable, and let $\{\omega_n : n \geq 0\}$ be an enumeration of its elements. Define $\omega \in \Omega$ so that $\omega(k) = \omega_k(k) + 1$ if $\omega_k(k) \neq D-1$ and $\omega(k) = \omega_k(k) - 1$ if $\omega_k(k) = D-1$. Show that $\omega \notin \{\omega_n : n \geq 1\}$ and therefore that there is no enumeration of Ω. In conjunction with (**i**), this proves that $\mathbb{R}$ is uncountable.

(iii) Let $n \in \mathbb{Z}$ and $\omega \in \tilde{\Omega}$ be given, and set $x = \sum_{k=0}^{\infty} \omega(k) D^{n-k}$. Show that $x \in \mathbb{Z}^+$ if and only if $n \geq 0$ and $\omega(k) = 0$ for all $k > n$. Next, say the ω is eventually periodic if there exists a $k_0 \geq 0$ and an $\ell \in \mathbb{Z}^+$ such that

$$\big(\omega(k_0 + m\ell + 1), \dots, \omega(k_0 + m\ell + \ell)\big) = \big(\omega(k_0 + 1), \dots, \omega(k_0 + \ell)\big)$$

for all $m \geq 1$. Show that $x > 0$ is rational if and only if ω is eventually periodic. The way to do the "only if" part is to write $x = \frac{a}{b}$, where $a \in \mathbb{N}$ and $b \in \mathbb{Z}^+$, and observe that is suffices to handle the case when $a < b$. Now think hard about how long division works. Specifically, at each stage either there is nothing to do or one of at most b possible numbers is be divided by b. Thus, either the process terminates after finitely many steps or, by at most the bth step, the number that is being divided by b is the same as one of the numbers that was divided by b at an earlier step.

Exercise 1.8 Here is a example, introduced by Cantor, of a set $C \subseteq [0, 1]$ that is uncountable but has an empty interior. Let $C_0 = [0, 1]$,

$$C_1 = [0, 1] \setminus \big(\tfrac{1}{3}, \tfrac{2}{3}\big) = \big[0, \tfrac{1}{3}\big] \cup \big[\tfrac{2}{3}, 1\big],$$

and, more generally, C_n is the union of the 2^n closed intervals that are obtained by removing the open middle third from each of the 2^{n-1} of which C_{n-1} consists. The set $C \equiv \bigcap_{n=0}^{\infty} C_m$ is called the *Cantor set*. Show that C is closed and $\text{int}(C) = \emptyset$. Further, take Ω to be the set of maps $\omega : \mathbb{N} \longrightarrow \{0, 2\}$, and Ω_1 equal to the set of $\omega \in \Omega$ such that $\omega(m) = 0$ for infinitely many $m \in \mathbb{N}$. Show there the map $\omega \rightsquigarrow \sum_{m=0} \omega(m)3^{-m-1}$ is one-to-one from Ω_1 onto $[0, 1) \cap \bigcap_{n=0}^{\infty} \text{int}(C_n)$, and use this to conclude that C is uncountable.

Exercise 1.9 Given $\mathbb{R}$-valued continuous functions f and g on a non-empty set $S \subseteq \mathbb{R}$, show that fg and, for all $\alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g$ are continuous. In addition, assuming that g never vanishes, show that $\frac{f}{g}$ is continuous. Finally, if f is a continuous function on $\emptyset \neq S_1 \subseteq \mathbb{R}$ with values in $S_2 \subseteq \mathbb{R}$ and if $g : S_2 \longrightarrow \mathbb{R}$ is continuous, show that their composition $g \circ f$ is again continuous.

Exercise 1.10 Suppose that a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ satisfies

$$f(x + y) = f(x) + f(y) \quad \text{for all } x, y \in \mathbb{R}. \tag{1.10.2}$$

Obviously, this equation will hold if $f(x) = f(1)x$ for all $x \in \mathbb{R}$. Cauchy asked under what conditions the converse is true, and for this reason (1.10.2) is sometimes called *Cauchy's equation*. Although the converse is known to hold in greater generality, the goal of this exercise is to show that the converse holds if f is continuous.

(i) Show that $f(mx) = mf(x)$ for any $m \in \mathbb{Z}$ and $x \in \mathbb{R}$, and use this to show that $f\left(\frac{x}{n}\right) = \frac{1}{n} f(x)$ for any $n \in \mathbb{Z} \setminus \{0\}$ and $x \in \mathbb{R}$.

(ii) From (i) conclude that $f\left(\frac{m}{n}\right) = \frac{m}{n} f(1)$ for all $m \in \mathbb{Z}$ and $n \in \mathbb{Z} \setminus \{0\}$, and then use continuity to show that $f(x) = f(1)x$ for all $x \in \mathbb{R}$.

Exercise 1.11 Let f be a function on one non-empty set S_1 with values in another non-empty set S_2. Show that if $\{A_\alpha : \alpha \in \mathcal{I}\}$ is any collection of subsets of S_2, then

$$f^{-1}\left(\bigcup_{\alpha \in \mathcal{I}} A_\alpha\right) = \bigcup_{\alpha \in \mathcal{I}} f^{-1}(A_\alpha) \text{ and } f^{-1}\left(\bigcap_{\alpha \in \mathcal{I}} A_\alpha\right) = \bigcap_{\alpha \in \mathcal{I}} f^{-1}(A_\alpha).$$

In addition, show that $f^{-1}(B \setminus A) = f^{-1}(B) \setminus f^{-1}(A)$ for $A \subseteq B \subseteq S$. Finally, give an example that shows that the second and last of these properties will not hold in general if f^{-1} is replaced by f.

Exercise 1.12 Show that $\tan = \frac{\sin}{\cos}$ is differentiable on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and that $\tan' = 1 + \tan^2 = \frac{1}{\cos^2}$ there. Use this to show that tan is strictly increasing on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, and let arctan denote the inverse of tan there. Finally, show that $\arctan' x = \frac{1}{1+x^2}$ for $x \in \mathbb{R}$.

Exercise 1.13 Show that

$$\lim_{x \to 0} \left(\frac{\sin x}{x}\right)^{\frac{1}{1-\cos x}} = e^{-\frac{1}{3}}.$$

In doing this computation, you might begin by observing that it suffices to show that

$$\lim_{x\to 0} \frac{1}{1-\cos x} \log\left(\frac{\sin x}{x}\right) = -\frac{1}{3}.$$

At this point one can apply L'Hôpital's rule, although it is probably easier to use Taylor's Theorem.

Exercise 1.14 Let $f : (a, b) \longrightarrow \mathbb{R}$ be a twice differentiable function. If $f'' \equiv f^{(2)}$ is continuous at the point $c \in (a, b)$, use Taylor's theorem to show that

$$f''(c) \equiv \partial^2 f(c) = \lim_{h\to\infty} \frac{f(c+h) + f(c-h) - 2f(c)}{h^2}.$$

Use this to show that if, in addition, f achieves its maximum value at c, then $f''(c) \leq 0$. Similarly, show that $f''(c) \geq 0$ if f achieves its minimum value at c. Hence, if f is twice continuously differentiable and one wants to locate the points at which f achieves its maximum (minimum) value, one need only look at points at which $f' = 0$ and $f'' \leq 0$ ($f'' \geq 0$). This conclusion is called the *second derivative test.*

Exercise 1.15 Suppose that $f : (a, b) \longrightarrow \mathbb{R}$ is differentiable at every $x \in (a, b)$. Darboux showed that f' has the intermediate value property. That is, for $a < c < d < b$ and y between $f'(c)$ and $f'(d)$, there is an $x \in [c, d]$ such that $f'(x) = y$. Darboux's idea is the following. Without loss in generality, assume that $f'(c) < y < f'(d)$, and consider the function $x \rightsquigarrow \varphi(x) \equiv f(x) - yx$. The function φ is continuous on $[c, d]$ and therefore achieves its minimum value at some $x \in [c, d]$. Show that $x \in (c, d)$ and therefore, by the first derivative test, that $0 = \varphi'(x) = f'(x) - y$. Finally, consider the function given by $f(x) = x^2 \sin \frac{1}{x}$ on $(-1, 1) \setminus \{0\}$ and 0 at 0, and show that f is differentiable at each $x \in (-1, 1)$ but that its derivative is discontinuous at 0.

Exercise 1.16 Let $f : (a, b) \longrightarrow \mathbb{R}$ be a differentiable function. Using (1.8.1), show that if f' is non-decreasing, then f is convex. In particular, if f is twice differentiable and $f'' \geq 0$, conclude that f is convex.

Exercise 1.17 Show that $-\log$ is convex on $(0, \infty)$, and use this to show that if $\{a_1, \ldots, a_n\} \subseteq (0, \infty)$ and $\{\theta_1, \ldots, \theta_n\} \subseteq [0, 1]$ with $\sum_{m=1}^{n} \theta_m = 1$, then

$$a_1^{\theta_1} \ldots a_n^{\theta_n} \leq \sum_{m=1}^{n} \theta_m a_m.$$

When $\theta_m = \frac{1}{n}$ for all m, this is the classical *arithmetic-geometric mean inequality.*

Exercise 1.18 Show that exp grows faster than any power of x in the sense that $\lim_{x\to\infty} \frac{x^\alpha}{e^x} = 0$ for all $\alpha > 0$. Use this to show that $\log x$ tends to infinity more slowly than any power of x in the sense that $\lim_{x\to\infty} \frac{\log x}{x^\alpha} = 0$ for all $\alpha > 0$. Finally, show that $\lim_{x\searrow 0} x^\alpha \log x = 0$ for all $\alpha > 0$.

Exercise 1.19 Show that $\prod_{m=1}^{\infty}\left(1-\frac{(-1)^n}{n}\right)$ converges but is not absolutely convergent.

Exercise 1.20 Just as is the case for absolutely convergent series (cf. Exercise 1.5), absolutely convergent products have the property that their products do not depend on the order in which their factors of multiplied. To see this, for a given $\epsilon > 0$, choose $m_\epsilon \in \mathbb{Z}^+$ so that

$$\left|\prod_{m=m_\epsilon+1}^{\infty}(1+|b_m|)-1\right| < \epsilon,$$

and conclude that if F is a finite subset of $\mathbb{Z}^+$ that contains $\{1,\dots,m_\epsilon\}$, then

$$\begin{aligned}\left|\prod_{m\in F}(1+b_m)-\prod_{m=1}^{\infty}(1+b_m)\right| &= \left|\prod_{m\in F}(1+b_m)\right|\left|1-\prod_{m\notin F}(1+b_m)\right|\\ &\le \prod_{m=1}^{\infty}(1+|b_m|)\left|1-\prod_{m=m_\epsilon+1}^{\infty}(1+|b_m|)\right| < \epsilon\prod_{m=1}^{\infty}(1+|b_m|).\end{aligned}$$

Use this to show that if $\{S_\ell : \ell \ge 1\}$ is an increasing sequence of subsets of $\mathbb{Z}^+$ and $\mathbb{Z}^+ = \bigcup_{\ell=1}^{\infty} S_\ell$, then $\lim_{\ell\to\infty}\prod_{m\in S_\ell}(1+b_m) = \prod_{m=1}^{\infty}(1+b_m)$.

Exercise 1.21 Show that every open subset G of $\mathbb{R}$ is the union of an at most countable number of mutually disjoint open intervals. To do this, for each rational numbers $r \in G$, let I_r be the set of $x \in G$ such that $[r,x] \subseteq G$ if $x \ge r$ or $[x,r] \subseteq G$ if $x \le r$. Show that I_r is an open interval and that either $I_r = I_{r'}$ or $I_r \cap I_{r'} = \emptyset$. Finally, let $\{r_n : n \ge 1\}$ be an enumeration of the rational numbers in G, choose $\{n_k : k \ge 1\}$ so that $n_1 = 1$ and, for $k \ge 2$, $n_k = \inf\{n > n_{k-1} : I_{r_n} \ne I_{r_m} \text{ for } 1 \le m \le n_{k-1}\}$. Show that $G = \bigcup_{k=1}^{K} I_{n_k}$ if $n_K < \infty = n_{K+1}$ and $G = \bigcup_{k=1}^{\infty} I_{n_k}$ if $n_k < \infty$ for all $k \ge 1$.

Chapter 2
Elements of Complex Analysis

It is frustrating that there is no $x \in \mathbb{R}$ for which $x^2 = -1$. In this chapter we will introduce a structure that provides an extension of the real numbers in which this equation has solutions.

2.1 The Complex Plane

To construct an extension of the real numbers in which solutions to $x^2 = -1$ exist, one has to go to the plane $\mathbb{R}^2$ and introduce a notion of addition and multiplication there. Namely, given (x_1, x_2) and (x_2, y_2), define their sum $(x_1, y_1) + (x_2, y_2)$ and product $(x_1, y_1)(x_2, y_2)$ to be

$$(x_1 + x_2, y_1 + y_2) \quad \text{and} \quad (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1).$$

Using the corresponding algebraic properties of the real numbers, one can easily check that

$$\begin{gathered}(x_1, y_1) + (x_2, y_2) = (x_2, y_2) + (x_1, y_1),\ (x_1, y_1)(x_2, y_2) = (x_2, y_2)(x_1, y_1),\\ \big((x_1, y_1)(x_2, y_2)\big)(x_3, y_3) = (x_1, y_1)\big((x_2, y_2)(x_3, y_3)\big),\\ \text{and } \big((x_1, y_1) + (x_2, y_2)\big)(x_3, y_3) = (x_1, y_1)(x_3, y_3) + (x_2, y_2)(x_3, y_3).\end{gathered}$$

Furthermore, $(0, 0)(x, y) = (0, 0)$ and $(1, 0)(x, y) = (x, y)$, and

$$\begin{gathered}(x_1, x_2) + (y_1, y_2) = (0, 0) \iff x_1 = -x_2 \ \&\ y_1 = -y_2,\\ (x_1, x_2) \neq (0, 0) \implies \Big[(x_1, x_2)(y_1, y_2) = (0, 0) \iff (y_1, y_2) = (0, 0)\Big].\end{gathered}$$

D.W. Stroock, *A Concise Introduction to Analysis*,
DOI 10.1007/978-3-319-24469-3_2

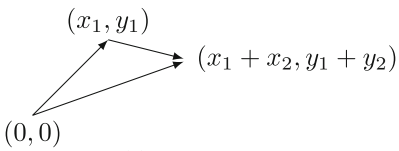
Thus, these operations on $\mathbb{R}^2$ satisfy all the basic properties of addition and multiplication on $\mathbb{R}$. In addition, when restricted to points of the form $(x, 0)$, they are given by addition and multiplication on $\mathbb{R}$:

$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0) \quad \text{and} \quad (x_1, 0)(x_2, 0) = (x_1 x_2, 0).$$

Hence, we can think of $\mathbb{R}$ with its additive and multiplicative structure as embedded by the map $x \in \mathbb{R} \longmapsto (x, 0) \in \mathbb{R}^2$ in $\mathbb{R}^2$ with its additive and multiplicative structure. The huge advantage of the structure on $\mathbb{R}^2$ is that the equation $(x, y)^2 = (-1, 0)$ has solutions, namely, (0, 1) and (0, −1) both solve it. In fact, they are the only solutions, since

$$(x, y)^2 = (-1, 0) \implies x^2 - y^2 = -1 \;\&\; 2xy = 0$$
$$\implies x^2 = -1 \;\&\; y = 0 \text{ or } x = 0 \;\&\; y^2 = 1 \implies x = 0 \;\&\; y = \pm 1.$$

The geometric interpretation of addition in the plane is that of vector addition: one thinks of (x_1, y_1) and (x_2, y_2) as vectors (i.e., arrows) pointing from the origin (0, 0) to the points in $\mathbb{R}^2$ that they represent, and one obtains $(x_1 + x_2, y_1 + y_2)$ by translating, without rotation, the vector for (x_2, y_2) to the vector that begins at the point of the vector for (x_1, y_1).



vector addition

To develop a feeling for multiplication, first observe that $(r, 0)(x, y) = (rx, ry)$, and so multiplication of (x, y) by $(r, 0)$ when $r \geq 0$ simply rescales the length of the vector for (x, y) by a factor of r. To understand what happens in general, remember that any point (x, y) in the plane has a *polar* representation $(r \cos\theta, r \sin\theta)$, where r is the length $\sqrt{x^2 + y^2}$ of the vector for (x, y) and θ is the angle that vector makes with the positive horizontal axis. If $(x_1, y_1) = (r_1 \cos\theta_1, r_1 \sin\theta_1)$ and $(x_2, y_2) = (r_2 \cos\theta_2, r_2 \sin\theta_2)$, then, by (1.5.1),

$$\begin{aligned}
&(x_1, y_1)(x_2, y_2) = (r_1 \cos\theta_1, r_1 \sin\theta_1)(r_2 \cos\theta_2, r_2 \sin\theta_2)\\
&= (r_1 r_2 \cos\theta_1 \cos\theta_2 - r_1 r_2 \sin\theta_1 \sin\theta_2, r_1 r_2 \cos\theta_1 \sin\theta_2 + r_1 r_2 \sin\theta_1 \cos\theta_2)\\
&= \big(r_1 r_2 \cos(\theta_1 + \theta_2), r_1 r_2 \sin(\theta_1 + \theta_2)\big)\\
&= \big(r_1, 0)\big(r_2 \cos(\theta_1 + \theta_2), r_2 \sin(\theta_1 + \theta_2)\big).
\end{aligned}$$

Hence, multiplying (x_2, y_2) by (x_1, y_1) rotates the vector for (x_2, y_2) through the angle θ_1 and rescales the length of the rotated vector by the factor r_1. Using this representation, it is evident that if $(x, y) = (r \cos\theta, r \sin\theta) \neq (0, 0)$ and

$$(x', y') = \big(r^{-1}\cos(-\theta), r^{-1}\sin(-\theta)\big) = (r^{-1}\cos\theta, -r^{-1}\sin\theta),$$

then $(x', y')(x, y) = (1, 0)$. Equivalently, if $(x, y) \neq (0, 0)$ the multiplicative inverse of (x, y) is $\left(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}\right)$.

Since $\mathbb{R}$ along with its arithmetic structure can be identified with $\{(x, 0) : x\in\mathbb{R}\}$ and its arithmetic structure, it is conventional to use the notation x instead of $(x, 0) = x(1, 0)$. Hence, if we set $i = (0, 1)$, then, with this convention, $(x, y) = x + iy$. When we use this notation, it is natural to think of $(x, y) = x + iy$ as some sort of number z. Of course z is not a "real" number, since it is an element of $\mathbb{R}^2$, not $\mathbb{R}$, and for that reason it is called a *complex* number and the set $\mathbb{C}$ of all complex numbers with the additive and multiplicative structure that we have been discussing is called the *complex plane*. For obvious reasons, the x in $z = x + iy$ is called the *real part* of z, and, because i is not a real number and for a long time people did not know how to rationalize its existence, the y is called the *imaginary* part of z.

Given $z = x + iy$, we will use $\mathfrak{R}(z) = x$ and $\mathfrak{I}(z) = y$ to denote its real and imaginary parts, and $|z|$ will denote its length $\sqrt{x^2 + y^2}$ (i.e., the distance between (x, y) and the origin). Either directly or using the polar representation of z, one can check that $|z_1 z_2| = |z_1||z_2|$. In addition, because the sum of the lengths of two sides of a triangle is at least the length of the third, one has the *triangle inequality* $|z_1 + z_2| \leq |z_1| + |z_2|$. Finally, in many computations involving complex numbers it is useful to consider the *complex conjugate* $\bar{z} \equiv x - iy$ of a complex number $z = x + iy$. For instance, $\mathfrak{R}(z) = \frac{z+\bar{z}}{2}$, $\mathfrak{I}(z) = \frac{z-\bar{z}}{i2}$, and $|z|^2 = |\bar{z}|^2 = z\bar{z}$. In terms of the polar representation $z = r\big(\cos\theta + i\sin\theta\big)$, $\bar{z} = r\big(\cos\theta - i\sin\theta\big)$, and from this it is easy to check that $\overline{z\zeta} = \bar{z}\bar{\zeta}$. Using these considerations, one can give another proof of the triangle inequality. Namely,

$$|z + \zeta|^2 = (z + \zeta)(\bar{z} + \bar{\zeta}) = |z|^2 + 2\mathfrak{R}(z\bar{\zeta}) + |\zeta|^2 \leq |z|^2 + 2|z\bar{\zeta}| + |\zeta|^2 = (|z| + |\zeta|)^2.$$

Since $|z_2 - z_1|$ measures the distance between z_1 and z_2, it is only reasonable that we say that the sequence $\{z_n : n \geq 1\}$ converges to $z \in \mathbb{C}$ if, for each $\epsilon > 0$, there exists an n_ϵ such that $|z - z_n| < \epsilon$ whenever $n \geq n_\epsilon$. Again, just as was the case for $\mathbb{R}$, there is a z to which $\{z_n : n \geq 1\}$ converges if and only if $\{z_n : n \geq 1\}$ satisfies Cauchy's criterion: for all $\epsilon > 0$ there is an n_ϵ such that $|z_n - z_m| < \epsilon$ whenever $n, m \geq n_\epsilon$. That is, $\mathbb{C}$ is complete. Indeed, the "only if" statement follows from the triangle inequality, since $|z_n - z_m| \leq |z_n - z| + |z - z_m| < \epsilon$ if $n, m \geq n_{\frac{\epsilon}{2}}$. To go the other direction, write $z_n = x_n + iy_n$ and note that $|x_n - x_m| \vee |y_n - y_m| \leq |z_n - z_m|$. Thus if $\{z_n : n \geq 1\}$ satisfies Cauchy's criterion, so do both $\{x_n : n \geq 1\}$ and $\{y_n : n \geq 1\}$. Hence, there exist $x, y \in \mathbb{R}$ such that $x_n \to x$ and $y_n \to y$. Now let $\epsilon > 0$ be given and choose n_ϵ so that $|x_n - x| \vee |y_n - y| < \frac{\epsilon}{\sqrt{2}}$ for $n \geq n_\epsilon$. Then $|z_n - z|^2 < \frac{\epsilon^2}{2} + \frac{\epsilon^2}{2} = \epsilon^2$ and therefore $|z_n - z| < \epsilon$ for $n \geq n_\epsilon$. The same sort of reduction allows one to use Theorem 1.3.3 to show that every bounded sequence $\{z_n : n \geq 1\}$ in $\mathbb{C}$ has a convergent subsequence.

All the results in Sect. 1.2 about series and Sect. 1.9 about products extend more or less immediately to the complex numbers. In particular, if $\{a_m : m \geq 1\} \subseteq \mathbb{C}$, then $\sum_{m=1}^{\infty} a_m$ converges if $\sum_{m=1}^{\infty} |a_m| < \infty$, in which case $\sum_{m=1}^{\infty} a_m$ is said to be *absolutely convergent*, and (cf. Exercise 1.5) the sum does not depend on the order in which the summands are added. Similarly, if $\{b_m : m \geq 1\} \subseteq \mathbb{C}$, then $\prod_{m=1}^{\infty}(1 + b_m)$ converges if $\sum_{m=1}^{\infty} |b_m| < \infty$, in which case $\prod_{m=1}^{\infty}(1+b_m)$ is said to be *absolutely convergent* and (cf. Exercise 1.20) the product does not depend on the order in which the factors are multiplied.

Just as in the case of $\mathbb{R}$, this notion of convergence has associated notions of open and closed sets. Namely, let $D(z, r) \equiv \{\zeta : |\zeta - z| < r\}$ be the open *disk*[1] of radius r centered at z, say that $G \subseteq \mathbb{C}$ is *open* if either $G = \emptyset$ or for each $z \in G$ there is an $r > 0$ such that $D(z, r) \subseteq G$, and say that $F \subseteq \mathbb{C}$ is *closed* if $F\complement$ is open. Further, given $S \subseteq \mathbb{C}$, the *interior* $\text{int}(S)$ of S is its largest open subset, and the *closure* $\bar{S}$ is its smallest closed superset. The connections between convergence and these concepts are examined in Exercises 2.1 and 2.4.

2.2 Functions of a Complex Variable

We next think about functions of a complex variable. In view of the preceding, we say that if $\emptyset \neq S \subseteq \mathbb{C}$ and $f : S \longrightarrow \mathbb{C}$, then f is *continuous* at a point $z \in S$ if for all $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(w) - f(z)| < \epsilon$ whenever $w \in S$ and $|w - z| < \delta$.

Using the triangle inequality, one sees that $\big||z_2| - |z_1|\big| \leq |z_2 - z_1|$, and therefore $z \in \mathbb{C} \longmapsto |z| \in [0, \infty)$ is continuous. In addition, by exactly the same argument as we used to prove Lemma 1.3.5, one can show that f is continuous at $z \in S$ if and only if $f(z_n) \longrightarrow f(z)$ whenever $\{z_n : n \geq 1\} \subseteq S$ tends to z. Thus, it is easy to check that linear combinations, products, and, if the denominator doesn't vanish, quotients of continuous functions are again continuous. In particular, polynomials of z with complex coefficients are continuous. Also, by the same arguments with which we proved Lemma 1.4.4 and Theorem 1.4.1, one can show that the uniform limit of continuous functions is again a continuous function and (cf. Exercise 2.4) that a continuous function on a bounded, closed set must be bounded and uniformly continuous.

We can now vastly increase the number of continuous functions of z that we know how to define by taking limits of polynomials. For this purpose, we introduce *power series*. That is, given $\{a_m : m \geq 0\} \subseteq \mathbb{C}$, consider the series $\sum_{m=0}^{\infty} a_m z^m$ on the set of $z \in \mathbb{C}$ for which it converges. Notice that if $z \in \mathbb{C}\setminus\{0\}$ and $\sum_{m=0}^{\infty} a_m z^m$ converges, then $\lim_{m\to\infty} a_m z^m = 0$, and so there is a $C < \infty$ such that $|a_m| \leq C|z|^{-m}$ for all m. Therefore, if $r = \frac{|\zeta|}{|z|}$, then $|a_m \zeta^m| \leq Cr^m$ for all $m \geq 0$. Hence, by the comparison

[1] Of course, thinking of $\mathbb{C}$ in terms of $\mathbb{R}^2$, if $z = x + iy$, then $D(z, r)$ is the ball of radius r centered at (x, y). However, to emphasize that we are thinking of $\mathbb{C}$ as the complex numbers, we will reserve the name disk and the notation $D(z, r)$ for balls in $\mathbb{C}$.

test, we see that if $\sum_{m=0}^{\infty} a_m z^m$ converges for some z, then $\sum_{m=0}^{\infty} a_m \zeta^m$ is absolutely convergent for all ζ with $|\zeta| < |z|$. As a consequence, we know that the interior of the set of z for which $\sum_{m=0}^{\infty} a_m z^m$ converges is an open disk centered at 0, and the radius of that disk is called the *radius of convergence* of $\sum_{m=0}^{\infty} a_m z^m$. Equivalently, the radius of convergence is the supremum of the set of $|z|$ for which $\sum_{m=0}^{\infty} a_m z^m$ converges.

Lemma 2.2.1 *Given a sequence $\{a_m : m \geq 0\} \subseteq \mathbb{C}$, the radius of convergence of $\sum_{m=0}^{\infty} a_m z^m$ is equal to the reciprocal of $\overline{\lim}_{m\to\infty} |a_m|^{\frac{1}{m}}$, when the reciprocal of 0 is interpreted as ∞ and that of ∞ is interpreted as 0. Furthermore, if $R \in (0, \infty)$ is strictly smaller than the radius of convergence, then there exists a $C_R < \infty$ and $\theta_R \in (0, 1)$ such that*

$$\sum_{m=n}^{\infty} |a_m||z|^m \leq C_R \theta_R^n \quad \text{for all } n \geq 0 \text{ and } |z| \leq R,$$

and so $\sum_{m=0}^{\infty} a_m z^m$ converges absolutely and uniformly on $\overline{D(0, R)}$ to a continuous function.

Proof First suppose that $\frac{1}{R} < \overline{\lim}_{m\to\infty} |a_m|^{\frac{1}{m}}$. Then $|a_m|R^m \geq 1$ for infinitely many m's, and so $a_m R^m$ does not tend to 0 and therefore $\sum_{m=0}^{\infty} a_m R^m$ cannot converge. Hence R is greater than or equal to the radius of convergence. Next suppose that $\frac{1}{R} > \overline{\lim}_{m\to\infty} |a_m|^{\frac{1}{m}}$. Then there is an m_0 such that $|a_m|R^m \leq 1$ for all $m \geq m_0$. Hence, if $|z| < R$ and $\theta = \frac{|z|}{R}$, then $|a_m z^m| \leq \theta^m$ for $m \geq m_0$, and so $\sum_{m=0}^{\infty} a_m z^m$ converges. Thus, in this case, R is less than or equal to the radius of convergence. Together these two imply the first assertion.

Finally, assume that $R > 0$ is strictly smaller that the radius of convergence. Then there exists an $r > R$ and $m_0 \in \mathbb{Z}^+$ such that $|a_m| \leq r^{-m}$ for all $m \geq m_0$, which means that there exists an $A < \infty$ such that $|a_m| \leq Ar^{-m}$ for all $m \geq 1$. Thus, if $\theta = \frac{R}{r}$, then

$$|z| \leq R \implies \sum_{m=n}^{\infty} |a_m z^m| \leq A\theta^n \sum_{m=0}^{\infty} \theta^m = \frac{A}{1-\theta}\theta^n. \qquad \square$$

We will now apply these considerations to extend to $\mathbb{C}$ some of the functions that we dealt with earlier, and we will begin with the function exp. Because $\frac{1}{m!} \geq 0$ and we already know that $\sum_{m=0}^{\infty} \frac{x^m}{m!}$ converges of all $x \in \mathbb{R}$, $\sum_{m=0}^{\infty} \frac{z^m}{m!}$ has an infinite radius of convergence. We can therefore define

$$\exp(z) = e^z = \sum_{m=0}^{\infty} \frac{z^m}{m!} \quad \text{for all } z \in \mathbb{C},$$

in which case we know that exp is a continuous function on $\mathbb{C}$. Because we obtained exp on $\mathbb{C}$ as a natural extension of exp on $\mathbb{R}$, it is reasonable to hope that some of the basic properties will extend as well. In particular, we would like to know that $e^{z_1+z_2} = e^{z_1}e^{z_2}$.

Lemma 2.2.2 *Given sequences* $\{a_n : n \ge 0\}$ *and* $\{b_n : n \ge 0\}$ *in* $\mathbb{C}$*, define* $c_n = \sum_{m=0}^{n} a_m b_{n-m}$*. Then the radii of convergence for* $\sum_{m=0}^{\infty}(a_m + b_m)z^m$ *and* $\sum_{m=0}^{\infty} c_m z^m$ *are at least as large as the smaller of the radii of convergence for* $\sum_{m=0}^{\infty} a_m z^m$ *and* $\sum_{m=0}^{\infty} b_m z^m$*. Moreover, if* $|z|$ *is strictly smaller than the radii of convergence for* $\sum_{m=0}^{\infty} a_m z^m$ *and* $\sum_{m=0}^{\infty} b_m z^m$*, then*

$$\sum_{m=0}^{\infty} a_m z^m + \sum_{m=0}^{\infty} b_m z^m = \sum_{m=0}^{\infty}(a_m + b_m)z^m$$

and

$$\left(\sum_{m=0}^{\infty} a_m z^m\right)\left(\sum_{m=0}^{\infty} b_m z^m\right) = \sum_{m=0}^{\infty} c_m z^m.$$

Proof Because verification of the assertions involving $\{a_n + b_n : n \ge 0\}$ is very easy, we will deal only with the assertions involving $\{c_n : n \ge 0\}$.

Suppose that $R > 0$ is strictly smaller than the radii of convergence of both $\sum_{m=0}^{\infty} a_m z^m$ and $\sum_{m=0}^{\infty} b_m z^m$. Then $|a_n| \vee |b_n| \le CR^{-n}$ for some $0 < C < \infty$, and so

$$|c_n| \le \sum_{m=0}^{n} |a_m||b_{n-m}| \le C^2(n+1)R^{-n}.$$

Hence $|c_n|^{\frac{1}{n}} \le C^{\frac{2}{n}}(n+1)^{\frac{1}{n}}R^{-1}$, and so, since (cf. Exercise 1.18) $\frac{1}{n}\log n \longrightarrow 0$, we know that the radius of convergence of $\sum_{m=0}^{\infty} c_m z^m$ is at least R and that the first assertion is proved. To prove the second assertion, let $z \ne 0$ of the specified sort be given. Then there exists an $R > |z|$ such that $|a_n| \vee |b_n| \le CR^{-n}$ for some $C < \infty$. Observe that, for each n,

$$\begin{aligned}\left(\sum_{m=0}^{n} a_m z^m\right)\left(\sum_{m=0}^{n} b_m z^m\right) &= \sum_{m=0}^{2n} z^m \sum_{\substack{0\le k,\ell\le n\\ k+\ell=m}} a_k b_\ell \\ &= \sum_{m=0}^{n} c_m z^m + \sum_{m=n+1}^{2n} z^m \sum_{k=m-n}^{n} a_k b_{m-k}.\end{aligned}$$

The product on the left tends to $\left(\sum_{m=0}^{\infty} a_m z^m\right)\left(\sum_{m=0}^{\infty} b_m z^m\right)$ as $n \to \infty$, and the first sum in the second line tends to $\sum_{m=0}^{\infty} c_m z^m$. Finally, if $\theta = \frac{|z|}{R}$, then the second term in the second line is dominated by

$$C^2 \sum_{m=n+1}^{2n} m\theta^m \le \frac{2C^2 n\theta^{n+1}}{1-\theta}.$$

Since $\theta \in (0,1)$, $\alpha \equiv -\log\theta > 0$, and so $n\theta^{n+1} \le \frac{n}{e^{\alpha n}} \le \frac{2n}{\alpha^2 n^2} \longrightarrow 0$ as $n \to \infty$. □

Set $a_n = \frac{z_1^n}{n!}$ and $b_n = \frac{z_2^n}{n!}$. By the same reasoning as we used to show that $\sum_{m=0}^{\infty} \frac{z^m}{m!}$ has an infinite radius of convergence, one sees that $\sum_{m=0}^{\infty} a_m z^m$ and $\sum_{m=0}^{\infty} b_m z^m$ do also. Next observe that

$$\sum_{m=0}^{n} a_m b_{n-m} = \sum_{m=0}^{n} \frac{z_1^m z_2^{n-m}}{m!(n-m)!} = \frac{1}{n!}\sum_{m=0}^{n}\binom{n}{m} z_1^m z_2^{n-m} = \frac{(z_1+z_2)^n}{n!},$$

and therefore, by Lemma 2.2.2 with $z=1$,

$$e^{z_1}e^{z_2} = \sum_{n=0}^{\infty}\frac{(z_1+z_2)^n}{n!} = e^{z_1+z_2}.$$

Our next goal is to understand what e^z really is, and, since $e^z = e^x e^{iy}$, the problem comes down to understanding e^z for purely imaginary z. For this purpose, note that if $\theta \in \mathbb{R}$, then

$$e^{i\theta} = \sum_{m=0}^{\infty} i^m \frac{\theta^m}{m!} = \sum_{m=0}^{\infty}(-1)^m\frac{\theta^{2m}}{(2m)!} + i\sum_{m=0}^{\infty}(-1)^m\frac{\theta^{2m+1}}{(2m+1)!},$$

which, by (1.8.5), means that

$$e^{i\theta} = \cos\theta + i\sin\theta \quad \text{for } \theta \in \mathbb{R}, \tag{2.2.1}$$

a remarkable equation known as *Euler's formula* in recognition of its discoverer. As a consequence, we know that when $z = x + iy$, then e^z corresponds to the point in $\mathbb{R}^2$ where the circle of radius e^x centered at the origin intersects the ray from the origin that passes through the point $(\cos y, \sin y)$.

With this information, we can now show that the equation $z^n = w$ has a solution for all $n \ge 1$ and $w \in \mathbb{C}$. Indeed, if $w = 0$ then $z = 0$ is the one and only solution. On the other hand, if $w \ne 0$, and we write $w = r(\cos\theta + i\sin\theta)$ where $\theta \in (-\pi, \pi]$, then $w = e^{\log r + i\theta}$, and so $z = e^{n^{-1}(\log r + i\theta)} = r^{\frac{1}{n}} e^{\frac{i\theta}{n}}$ satisfies $z^n = w$. However, this isn't the only solution. Namely, we could have written w as $\exp\bigl(\log r + i(\theta + 2m\pi)\bigr)$ for any $m \in \mathbb{Z}$, in which case we would have concluded that

$$z_m \equiv r^{\frac{1}{n}} \exp\bigl(in^{-1}(\theta + 2m\pi)\bigr)$$

is also a solution. Of course, $z_{m_1} = z_{m_2} \iff (m_2 - m_1)$ is divisible by n, and so we really have only n distinct solutions, $z_0, \dots, z_{n-1}$. That this list contains all the solutions follows from a simple algebraic lemma.

Lemma 2.2.3 *Suppose that* $n \geq 1$ *and* $f(z) = \sum_{m=0}^{n} a_m z^m$ *where* $a_n \neq 0$. *If* $f(\zeta) = 0$, *then* $f(z) = (z - \zeta)g(z)$, *where* $g(z) = \sum_{m=0}^{n-1} b_m z^m$ *for some choice of* $b_0, \dots, b_{n-1} \in \mathbb{C}$. *In particular, there can be no more than n distinct solutions to* $f(z) = 0$. *(See the fundamental theorem of algebra in Sect. 6.2 for more information).*

Proof If $f(0) = 0$, then $a_0 = 0$, and so $f(z) = z \sum_{m=0}^{n-1} a_{m+1} z^m$. Next suppose that $f(\zeta) = 0$ for some $\zeta \neq 0$, and consider $f_\zeta(z) = f(z+\zeta)$. By the binomial theorem, f_ζ is again an nth order polynomial, and clearly $f_\zeta(0) = 0$. Hence $f(z + \zeta) = z g_\zeta(z)$, where g_ζ is a polynomial of order $(n - 1)$, which means that we can take $g(z) = g_\zeta(z - \zeta)$.

Given the first part, the second part follows easily. Indeed, if $f(\zeta_m) = 0$ for distinct $\zeta_1, \dots, \zeta_{n+1}$, then, by repeated application of the first part, $f(z) = a_n \prod_{m=1}^{n} (z - \zeta_m)$. However, because $a_n \neq 0$ and $\zeta_{n+1} \neq \zeta_m$ for any $1 \leq m \leq n$, this means that $f(\zeta_{n+1})$ could not have been 0. □

We have now found all the solutions to $z^n = w$. When $w = 1$, the solutions are $e^{i\frac{2\pi m}{n}}$ for $0 \leq m < n$, and these are called the n *roots of unity*. Obviously, for any $w \neq 0$, all solutions can be generated from any particular solution by multiplying it by the roots of unity.

Another application of (2.2.1) is the extension of log to $\mathbb{C} \setminus \{0\}$. For this purpose, think of log on $(0, \infty)$ as the inverse of exp on $\mathbb{R}$. Then, (2.2.1) tells us how to extend log to $\mathbb{C} \setminus \{0\}$ as the inverse of exp on $\mathbb{C}$. Indeed, given $z \neq 0$, set $r = |z|$ and take θ to be the unique element of $(-\pi, \pi]$ for which $z = re^{i\theta}$. Then, if we take $\log z = \log r + i\theta$, it is obvious that $e^{\log z} = z$. To get a more analytic description, note that cos is a strictly decreasing function from $(0, \pi)$ onto $(-1, 1)$, define arccos on $(-1, 1)$ to be its inverse, and extend arccos as a continuous function on $[-1, 1]$ by taking $\arccos(-1) = \pi$ and $\arccos 1 = 0$. Then

$$\log z = \tfrac{1}{2}\log(x^2 + y^2) + i \begin{cases} \arccos \frac{x}{\sqrt{x^2+y^2}} & \text{if } y \geq 0 \\ -\arccos \frac{x}{\sqrt{x^2+y^2}} & \text{if } y < 0 \end{cases} \tag{2.2.2}$$

for $z = x + iy \neq 0$. Observe that, although log is defined on the whole of $\mathbb{C} \setminus \{0\}$, it is continuous only on $\mathbb{C} \setminus (-\infty, 0]$, where $(-\infty, 0]$ here denotes the subset of $\mathbb{C}$ corresponding to the subset $\{(x, 0) : x \leq 0\}$ of $\mathbb{R}^2$.

2.3 Analytic Functions

Given a non-empty, open set $G \subseteq \mathbb{C}$ and a function $f : G \longrightarrow \mathbb{C}$, say that f is *differentiable* in the direction θ at $z \in G$ if the limit

$$f'_\theta(z) \equiv \lim_{t\to 0} \frac{f(z+te^{i\theta}) - f(z)}{t} \quad \text{exists.}$$

If f is differentiable at $z \in G$ in every direction, it is said to be differentiable at z, and if f is differentiable at every $z \in G$ it is said to the differentiable on G. Further, if f'_θ is continuous at z for every θ, then we say that f is continuously differentiable at z, and when f is continuously differentiable at every $z \in G$, we say that it is continuously differentiable on G.

Say that an open set $G \subseteq \mathbb{C}$ is *connected* if it cannot be written as the union of two non-empty, disjoint open sets. (See Exercise 4.5 for more information.)

Lemma 2.3.1 *Let $z \in \mathbb{C}$, $R > 0$, and $f : D(z, R) \longrightarrow \mathbb{C}$ be given, and assume that f is differentiable in the directions 0 and $\frac{\pi}{2}$ at every point in $D(z, R)$. If both f'_0 and $f'_{\frac{\pi}{2}}$ are continuous at z, then f is continuously differentiable at z and*

$$f'_\theta(z) = f'_0(z)\cos\theta + f'_{\frac{\pi}{2}}(z)\sin\theta \quad \textit{for each } \theta \in (-\pi, \pi].$$

Furthermore, if f'_0 and $f'_{\frac{\pi}{2}}$ are continuous on $D(z, R)$, then f is continuously differentiable on $D(z, R)$ and

$$\lim_{t\to 0} \sup_{\theta\in(-\pi,\pi]} \left| \frac{f\big(z+te^{i\theta}\big) - f(z)}{t} - f'_\theta(z) \right| = 0.$$

Finally, if $G \subseteq \mathbb{C}$ is a connected open set and f is differentiable on G with $f'_\theta = 0$ on G for all $\theta \in (-\pi, \pi]$, then f is constant on G.

Proof First observe that f determines functions $u : D(z, R) \longrightarrow \mathbb{R}$ and $v : D(z, R) \longrightarrow \mathbb{R}$ such that $f = u + iv$ and that, for any $\theta \in (-\pi, \pi]$, $f'_\theta(\zeta)$ exists at a point $\zeta \in D(z, R)$ if and only if both $u'_\theta(\zeta)$ and $v'_\theta(\zeta)$ do, in which case $f'_\theta(\zeta) = u'_\theta(\zeta) + iv'_\theta(\zeta)$. Furthermore, if $\zeta \in D(z, R)$ and f'_θ exists on $D(\zeta, r)$ for some $0 < r < R - |\zeta - z|$, then f'_θ is continuous at ζ if and only if both u'_θ and v'_θ are. Thus, without loss in generality, from now on we may and will assume that f is $\mathbb{R}$-valued.

First assume that f'_0 and $f'_{\frac{\pi}{2}}$ are continuous at z. Then, by Theorem 1.8.1, for $0 < |t| < R$,

$$\begin{aligned} &f(z+te^{i\theta}) - f(z) \\ &\quad = f(z + t\cos\theta + it\sin\theta) - f(z + it\sin\theta) + f(z + it\sin\theta) - f(z) \\ &\quad = tf'_0(z + \tau_1\cos\theta + it\sin\theta)\cos\theta + tf'_{\frac{\pi}{2}}(z + i\tau_2\sin\theta)\sin\theta \end{aligned}$$

for some choice of τ_1 and τ_2 with $|\tau_1| \vee |\tau_2| < |t|$. Hence, after dividing through by t and then letting $t \to 0$, we see that $f'_\theta(z)$ exists and is equal to $f'_0(z)\cos\theta + f'_{\frac{\pi}{2}}(z)\sin\theta$.

Next assume that f_0' and $f_{\frac{\pi}{2}}'$ are continuous on $D(z,R)$. Then, by Theorem 1.8.1 and the preceding,

$$\begin{aligned} &f(z+te^{i\theta}) - f(z) - tf_\theta'(z) \\ &\quad = t\big(f_0'(z+\tau e^{i\theta}) - f_0'(z)\big)\cos\theta + t\big(f_{\frac{\pi}{2}}'(z+\tau e^{i\theta}) - f_{\frac{\pi}{2}}'(z)\big)\sin\theta \end{aligned}$$

for some $0 < |\tau| < t$. Hence, if for a given $\epsilon > 0$ we choose $\delta > 0$ so that $|f_0'(z+\zeta) - f_0'(z)| \vee |f_{\frac{\pi}{2}}'(z+\zeta) - f_{\frac{\pi}{2}}'(z)| < \frac{\epsilon}{2}$ when $|\zeta| < \delta$, then

$$\sup_{\theta\in(-\pi,\pi]} \left| \frac{f(z+te^{i\theta}) - f(z)}{t} - f_\theta'(z) \right| < \epsilon \quad \text{for all } t \in (-\delta,\delta)\setminus\{0\}.$$

Finally, assume that, for each θ, f_θ' exists and is equal to 0 on a connected open set G, let $\zeta \in G$, and set $c = f(\zeta)$. Obviously, $G_1 \equiv \{z \in G : f(z) \neq c\}$ is open. Next suppose $f(z) = c$, and choose $R > 0$ so that $\overline{D(z,R)} \subseteq G$. Given $z' \in D(z,R)\setminus\{z\}$, express $z' - z$ as $re^{i\theta}$, apply (1.8.1) to see that $f(z') - f(z) = f_\theta'(z + \rho e^{i\theta})r$ for some $0 < |\rho| < r$, and conclude that $f(z') = c$. Hence, $G_2 \equiv \{z \in G : f(z) = c\}$ is also open, and therefore, since $G = G_1 \cup G_2$ and $\zeta \in G_2$, $G_1 = \emptyset$. □

We now turn our attention to a very special class of differentiable functions f on $\mathbb{C}$, those that depend on z thought of as a number rather than as a vector in $\mathbb{R}^2$. To be more precise, we will be looking at f's with the property that the amount $f(z+\zeta) - f(z)$ by which f changes when z is displaced by a sufficiently small ζ is, to first order, given by a number $f'(z) \in \mathbb{C}$ times ζ. That is, if f is a $\mathbb{C}$-valued function on an open set $G \subseteq \mathbb{C}$ and $z \in G$, then

$$f'(z) \equiv \lim_{\zeta\to 0} \frac{f(z+\zeta) - f(z)}{\zeta} \quad \text{exists.}$$

Further we will assume that f' is continuous on G. Such functions are said to be *analytic* on G. It is important to recognize that most differentiable functions are not analytic. Indeed, a continuously differentiable function on G is analytic there if and only if

$$f_\theta'(z) = e^{i\theta} f_0'(z) \quad \text{for all } z \in G \text{ and } \theta \in (-\pi,\pi].$$

To see this, first suppose that f is continuously differentiable on G and that the preceding holds. Then, by Lemma 2.3.1, for any $z \in G$

$$\lim_{r\searrow 0} \sup_{\theta\in(-\pi,\pi]} \left| \frac{f(z+re^{i\theta}) - f(z)}{r} - e^{i\theta} f_0'(z) \right| = 0,$$

and so, after dividing through by $e^{i\theta}$, one sees that f is analytic and that $f'(z) = f_0'(z)$. Conversely, suppose that f is analytic on G. Then

$$e^{-i\theta} f_\theta'(z) = \lim_{r \searrow 0} \frac{f(z + re^{i\theta}) - f(z)}{re^{i\theta}} = f'(z),$$

and so $f_\theta'(z) = e^{i\theta} f'(z) = e^{i\theta} f_0'(z)$. Writing $f = u + iv$, where u and v are $\mathbb{R}$-valued, and taking $\theta = \frac{\pi}{2}$, one sees that, for analytic f's, $u_{\frac{\pi}{2}}' + iv_{\frac{\pi}{2}}' = e^{i\frac{\pi}{2}}(u_0' + iv_0') = iu_0' - v_0'$ and therefore

$$u_0' = v_{\frac{\pi}{2}}' \text{ and } u_{\frac{\pi}{2}}' = -v_0'. \tag{2.3.1}$$

Conversely, if $f = u + iv$ is continuously differentiable on G and (2.3.1) holds, then, by Lemma 2.3.1

$$\begin{aligned} f_\theta' &= f_0' \cos\theta + f_{\frac{\pi}{2}}' \sin\theta = u_0' \cos\theta + iv_0' \cos\theta + u_{\frac{\pi}{2}}' \sin\theta + iv_{\frac{\pi}{2}}' \sin\theta \\ &= (\cos\theta + i \sin\theta)u_0' + i(\cos\theta + i\sin\theta)v_0' = e^{i\theta} f_0', \end{aligned}$$

and so f is analytic. The equations in (2.3.1) are called the *Cauchy–Riemann equations*, and using them one sees why so few functions are analytic. For instance, (2.3.1) together with the last part of Lemma 2.3.1 show that the only $\mathbb{R}$-valued analytic functions on a connected open set are constant.

The obvious challenge now is that of producing interesting examples of analytic functions. Without any change in the argument used in the real-valued case, it is easy to see that if $f(z) = \sum_{m=0}^n a_m z^m$, then f is analytic on $\mathbb{C}$ and $f'(z) = \sum_{m=0}^{n-1}(m+1)a_{m+1}z^m$. As the next theorem shows, the same is true for power series.

Theorem 2.3.2 *Given $\{a_m : m \geq 0\} \subseteq \mathbb{C}$ and a function $\psi : \mathbb{N} \longrightarrow \mathbb{C}$ satisfying $1 \leq |\psi(m)| \leq Cm^k$ for some $C < \infty$ and $k \in \mathbb{N}$, $\sum_{n=0}^\infty \psi(n)a_n z^n$ has the same radius of convergence as $\sum_{m=0}^\infty a_m z^m$. Moreover, if $R > 0$ is strictly less than the radius of convergence of $\sum_{m=0}^\infty a_m z^m$ and $f(z) = \sum_{m=0}^\infty a_m z^m$ for $z \in D(0, R)$, then f is analytic on $D(0, R)$ and $f'(z) = \sum_{m=0}^\infty (m+1)a_{m+1}z^m$.*

Proof Set $b_m = \psi(m)a_m$. Obviously $\overline{\lim}_{m\to\infty} |b_m|^{\frac{1}{m}} \geq \overline{\lim}_{m\to\infty} |a_m|^{\frac{1}{m}}$. At the same time, $|b_m|^{\frac{1}{m}} \leq C^{\frac{1}{m}} m^{\frac{k}{m}} |a_m|^{\frac{1}{m}}$ and, by Exercise 1.18,

$$\log(C^{\frac{1}{m}} m^{\frac{k}{m}}) = \frac{1}{m}\big(\log C + k \log m\big) \longrightarrow 0 \quad \text{as } m \to \infty.$$

Hence, $\overline{\lim}_{m\to\infty} |b_m|^{\frac{1}{m}} \leq \overline{\lim}_{m\to\infty} |a_m|^{\frac{1}{m}}$. By Lemma 2.2.1, this completes the proof of the first assertion.

Turning to the second assertion, let $z,\ \zeta \in D(0,r)$ for some $0 < r < R$. Then

$$f(\zeta) - f(z) = \sum_{m=0}^{\infty} a_m(\zeta^m - z^m) = (\zeta - z)\sum_{m=1}^{\infty} a_m \sum_{k=1}^{m} \zeta^{k-1} z^{m-k}$$
$$= (\zeta - z)\sum_{m=1}^{\infty} m a_m z^{m-1} + (\zeta - z)\sum_{m=1}^{\infty} a_m \sum_{k=1}^{m}\big(\zeta^{k-1}z^{m-k} - z^{m-1}\big).$$

Hence, what remains is to show that $\sum_{m=1}^{\infty} a_m \sum_{k=1}^{m}\big(\zeta^{k-1}z^{m-k} - z^{m-1}\big) \longrightarrow 0$ as $\zeta \to z$. But there exists a $C < \infty$ such that, for any M,

$$\left|\sum_{m=1}^{\infty} a_m \sum_{k=1}^{m}\big(\zeta^{k-1}z^{m-k} - z^{m-1}\big)\right|$$
$$\leq \sum_{m=0}^{M} |a_m| \sum_{k=1}^{m} |\zeta^{k-1}z^{m-k} - z^{m-1}| + C \sum_{m=M+1}^{\infty} m\left(\frac{r}{R}\right)^{m-1}.$$

Since, for each M, the first term on the right tends to 0 as $\zeta \to z$, it suffices to show that the second term tends to 0 as $M \to \infty$. But this follows immediately from the first part of this lemma and the observation that $\frac{r}{R}$ is strictly smaller than the radius of convergence of the geometric series. □

As a consequence of Theorem 2.3.2, we know that

$$\exp(z) = \sum_{m=0}^{\infty} \frac{z^m}{m!},\quad \sin z \equiv \sum_{m=0}^{\infty} (-1)^m \frac{z^{2m+1}}{(2m+1)!},\quad \text{and}\quad \cos z \equiv \sum_{m=0}^{\infty} (-1)^m \frac{z^{2m}}{(2m)!}$$

are all analytic functions on the whole of $\mathbb{C}$. In addition, it remains true that $e^{iz} = \cos z + i \sin z$ for all $z \in \mathbb{C}$. Therefore, since $\sin(-z) = -\sin z$ and $\cos(-z) = \cos z$, $\cos z = \frac{e^{iz}+e^{-iz}}{2}$ and $\sin z = \frac{e^{iz}-e^{-iz}}{2i}$. The functions $\sinh x \equiv -i\,\sin(ix) = \frac{e^x - e^{-x}}{2}$ and $\cosh x \equiv \cos(ix) = \frac{e^x+e^{-x}}{2}$ for $x \in \mathbb{R}$ are called the *hyperbolic* sine and cosine functions.

Notice that when $f(z) = \sum_{m=0}^{\infty} a_m z^m$ as in Theorem 2.3.2, not only f but also all its derivatives are analytic in $D(0,R)$. Indeed, for $k \geq 1$, the kth derivative $f^{(k)}$ will be given by

$$\sum_{m=k}^{\infty} m(m-1)\cdots(m-k+1)a_m z^{m-k},$$

and so $a_m = \frac{f^{(m)}(0)}{m!}$. Hence, the series representation of this f can be thought of as a Taylor series. It turns out that this is no accident. As we will show in Sect. 6.2 (cf. Theorem 6.2.2), every analytic function f in a disk $D(\zeta, R)$ has derivatives of all orders, the radius of convergence of $\sum_{m=0}^{\infty} \frac{f^{(n)}(\zeta)}{n!} z^n$ is at least R, and $f(z) = \sum_{m=0}^{\infty} \frac{f^{(m)}(\zeta)}{m!}(z-\zeta)^m$ for all $z \in D(\zeta, R)$.

The function log on $G = \mathbb{C} \setminus (-\infty, 0]$ provides a revealing example of what is going on here. By (2.2.2), $\log z = u(z) + iv(z)$, where $u(z) = \frac{1}{2}\log(x^2+y^2)$ and $v(z) = \pm \arccos \frac{x}{\sqrt{x^2+y^2}}$, the sign being the same as that of y. By a straightforward application of the chain rule, one can show that $u_0'(z) = \frac{x}{x^2+y^2}$ and $u_{\frac{\pi}{2}}'(z) = \frac{y}{x^2+y^2}$. Before computing the derivatives of v, we have to know how to differentiate arccos. Since $-\arccos$ is strictly increasing on $(0, \pi)$, Theorem 1.8.4 applies and says that $-\arccos' \xi = \frac{1}{\sin(\arccos \xi)}$ for $\xi \in (-1, 1)$. Next observe that $\sin = \pm\sqrt{1-\cos^2}$ and therefore, because $\sin > 0$ on $(0, \pi)$, $\arccos' \xi = -\frac{1}{\sqrt{1-\xi^2}}$ for $\xi \in (-1, 1)$. Applying the chain rule again, one sees that $v_0'(z) = -\frac{y}{x^2+y^2}$ and $v_{\frac{\pi}{2}}'(z) = \frac{x}{x^2+y^2}$, first for $y \neq 0$ and then, by Lemma 1.8.3, for all $y \in \mathbb{R}$. Hence, u and v satisfy the Cauchy–Riemann equations and so log is analytic on G. In fact,

$$\log' z = u_0'(z) + iv_0'(z) = \frac{x - iy}{x^2+y^2} = \frac{\bar{z}}{|z|^2},$$

and so

$$\log' z = \frac{1}{z} \quad \text{for } z \in \mathbb{C} \setminus (-\infty, 0].$$

With this information, we can now show that

$$\log \zeta = -\sum_{m=1}^{\infty} \frac{(1-\zeta)^m}{m} \quad \text{for } \zeta \in D(1, 1). \tag{2.3.2}$$

To check this, note that the derivative of the right hand side is equal to

$$\sum_{m=1}^{\infty} (-1)^{m-1}(\zeta - 1)^{m-1} = \sum_{m=0}^{\infty} (1-\zeta)^m = \frac{1}{\zeta}.$$

Hence, the difference g between the two sides of (2.3.2) is an analytic function on $D(1, 1)$ whose derivative vanishes. Since $g_\theta' = e^{i\theta} g'$ and (cf. Exercise 4.5) G is connected, the last part of Lemma 2.3.1 implies that g is constant, and so, since $g(1) = 0$, (2.3.2) follows. Given (2.3.2), one can now say that if $z \in G$ and $R = \inf\{|\xi - z| : \xi \le 0\}$, then

$$\log \zeta = \log z - \sum_{m=1}^{\infty} \frac{(z-\zeta)^m}{m z^m} \quad \text{for } \zeta \in D(z, R).$$

Indeed,

$$\zeta = z\frac{\zeta}{z} = e^{\log z} e^{\log \frac{\zeta}{z}} = e^{\log z + \log \frac{\zeta}{z}},$$

from which we conclude that $\varphi(\zeta) \equiv \frac{\log\zeta - \log z - \log\frac{\zeta}{z}}{i2\pi}$ must be an integer for each $\zeta \in D(z, R)$. But φ is a continuous function on $D(z, R)$ and $\varphi(z) = 0$. Thus if $\varphi(\zeta) \neq 0$ for some $\zeta \in D(z, R)$ and if $u(t) = \varphi\big((1-t)z + t\zeta\big)$, then u would be a continuous, integer-valued function on $[0, 1]$ with $u(0) = 0$ and $|u(1)| \geq 1$, which, by Theorem 1.3.6, would lead to the contradiction that $|u(t)| = \frac{1}{2}$ for some $t \in (0, 1)$. Therefore, we know that $\log\zeta = \log z + \log\frac{\zeta}{z}$ and, since $|\zeta - z| < |z|$ and therefore $\left|\frac{\zeta}{z} - 1\right| < 1$, we can apply (2.3.2) to complete the calculation.

2.4 Exercises

Exercise 2.1 Many of the results about open sets, convergence, and continuity for $\mathbb{R}$ have easily proved analogs for $\mathbb{C}$, and these are the topic of this and Exercise 2.4. Show that F is closed if and only if $z \in F$ whenever there is a sequence in F that converges to z. Next, given $S \subseteq \mathbb{C}$, show that $z \in \text{int}(S)$ if and only if $D(z, r) \subseteq S$ for some $r > 0$, and show that $z \in \bar{S}$ if and only if there is a sequence in S that converges to z. Finally, define the notion of subsequence in the same way as we did for $\mathbb{R}$, and show that if K is a bounded (i.e., there is an M such that $|z| \leq M$ for all $z \in K$), closed set, then every sequence in K admits a subsequence that converges to some point in K.

Exercise 2.2 Let $R > 0$ be given, and consider the circles $\mathbb{S}^1(0, R)$ and $\mathbb{S}^1(2R, R)$. These two circles touch at the point $\zeta(0) = (R, 0)$. Now imagine rolling $\mathbb{S}^1(2R, R)$ counter clockwise along $\mathbb{S}^1(0, R)$ for a distance θ so that it becomes the circle $\mathbb{S}^1(2Re^{i\theta}, R)$, and let $\zeta(\theta)$ denote the point to which $\zeta(0)$ is moved. Show that $\zeta(\theta) = R\big(2e^{i\theta} - e^{i2\theta}\big)$. Because it is somewhat heart shaped, the path $\{\zeta(\theta) : \theta \in [0, 2\pi)\}$ is called a *cardiod*. Next set $z(\theta) = \zeta(\theta) - R$, and show that $z(\theta) = 2R(1 - \cos\theta)e^{i\theta}$.

Exercise 2.3 Given sequences $\{a_n : n \geq 1\}$ and $\{b_n : n \geq 1\}$ of complex numbers, prove *Schwarz's inequality*

$$\sum_{m=0}^{\infty} |a_m||b_m| \leq \sqrt{\sum_{m=0}^{\infty} |a_m|^2}\sqrt{\sum_{m=0}^{\infty} |b_m|^2}.$$

Clearly, it suffices to treat the case when all the a_n's and b_n's are non-negative and there is an N such that $a_n = b_n = 0$ for $n > N$ and $\sum_{n=1}^{N} a_n^2 > 0$. In that case consider the function

$$\varphi(t) = \sum_{n=1}^{N}(a_n t + b_n)^2 = t^2\sum_{n=1}^{N} a_n^2 + 2t\sum_{n=1}^{N} a_n b_n + \sum_{n=1}^{N} b_n^2 \quad \text{for } t \in \mathbb{R}.$$

Observe that $\varphi(t)$ is non-negative for all t and that

$$\varphi(t_0) = -\frac{\left(\sum_{n=1} a_n b_n\right)^2}{\sum_{n=1}^N a_n^2} + \sum_{n=1}^N b_n^2 \quad \text{if } t_0 = -\frac{\sum_{n=1}^N a_n b_n}{\sum_{n=1}^N a_n^2}.$$

Exercise 2.4 If $H \neq \emptyset$ is open and $f : H \longrightarrow \mathbb{C}$, show that f is continuous if and only if $f^{-1}(G)$ is open whenever G is. If K is closed and bounded and if $f : K \longrightarrow \mathbb{C}$ is continuous, show that f is bounded (i.e., $\sup_{z\in K} |f(z)| < \infty$) and is uniformly continuous in the sense that for each $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(\zeta) - f(z)| < \epsilon$ whenever $z, \zeta \in K$ with $|\zeta - z| < \delta$.

Exercise 2.5 Here is an interesting, totally non-geometric characterization of the trigonometric functions. Thus, in doing this exercise, forget about the geometric description of the sine and cosine functions.

(i) Using Taylor's theorem, show that for each $a, b \in \mathbb{R}$ the one and only twice differentiable $f : \mathbb{R} \longrightarrow \mathbb{R}$ that satisfies $f'' = -f$, $f(0) = a$, and $f'(0) = b$ is the function $f(x) = ac(x) + bs(x)$, where

$$c(x) \equiv \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m}}{(2m)!} \quad \text{and} \quad s(x) \equiv \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m+1}}{(2m+1)!}.$$

(ii) Show that $c' = -s$ and $s' = c$, and conclude that $c^2 + s^2 = 1$.

(iii) Using (i) and (ii), show that

$$c(x + y) = c(x)c(y) - s(x)s(y) \quad \text{and} \quad s(x + y) = s(x)c(y) + s(y)c(x).$$

(iv) Show that $c(x) \geq \frac{1}{2}$ and therefore that $s(x) \geq \frac{x}{2}$ for $x \in \left[0, \frac{1}{2}\right]$. Thus, if $L > \frac{1}{2}$ and $c \geq 0$ on $[0, L]$, then $s(x) \geq \frac{1}{4}$ on $\left[\frac{1}{2}, L\right]$, and so

$$0 \leq c(L) \leq c\left(\tfrac{1}{2}\right) - \tfrac{1}{4}\left(L - \tfrac{1}{2}\right) \leq 1 - \tfrac{2L-1}{8}.$$

Therefore $L \leq \frac{9}{2}$, which means that there is an $\alpha \in \left(\frac{1}{2}, \frac{9}{2}\right]$ such that $c(\alpha) = 0$ and $c(x) > 0$ for $x \in [0, \alpha)$. Show that $s(\alpha) = 1$.

(v) By combining (ii) and (iv), show that $c(\alpha + x) = -s(x)$ and $s(\alpha + x) = c(x)$, and conclude that

$$\begin{array}{lll} c(2\alpha + x) = -s(x) & s(2\alpha + x) = -c(x) & c(3\alpha + x) = s(x) \\ s(3\alpha + x) = -c(x) & c(4\alpha + x) = c(x) & s(4\alpha + x) = s(x). \end{array}$$

In particular, this shows that c and s are periodic with period 4α.

Clearly, by uniqueness, $c(x) = \cos x$ and $s(x) = \sin x$, where cos and sin are defined geometrically in terms of the unit circle. Thus, $\alpha = \frac{\pi}{2}$, one fourth the circumference of the unit disk. See the discussion preceding Corollary 3.2.4 for a more satisfying treatment of this point.

Exercise 2.6 Given $\omega \in \mathbb{C} \setminus \{0\}$, define

$$\sin_\omega x = \frac{e^{\omega x} - e^{-\omega x}}{2\omega} \quad \text{and} \quad \cos_\omega x = \frac{e^{\omega x} + e^{-\omega x}}{2}$$

for $x \in \mathbb{R}$. Obviously, sin and cos are $\sin_i$ and $\cos_i$, and sinh and cosh are $\sin_1$ and $\cos_1$. Show that $\sin_\omega$ and $\cos_\omega$ satisfy the equation $u''_\omega = \omega^2 u$ on $\mathbb{R}$ and that if u is any solution to this equation then $u(x) = u(0)\cos_\omega x + u'(0)\sin_\omega x$.

Exercise 2.7 Show that if f and g are analytic functions on an open set G, then $\alpha f + \beta g$ is analytic for all $\alpha, \beta \in \mathbb{C}$ as is fg. In fact, show that $(\alpha f + \beta g)' = \alpha f' + \beta g'$ and $(fg)' = f'g + fg'$. In addition, if g never vanishes, show that $\frac{f}{g}$ is analytic and $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$. Finally, show that if f is an analytic function on an open set G_1 with values in an open set G_2 and if g is an analytic function on G_2, then their composition $g \circ f$ is analytic on G_1 and $(g \circ f)' = (g' \circ f)f'$. Hence, for differentiation of analytic functions, linearity as well as the product, quotient, and chain rules hold.

Exercise 2.8 Given $\omega \in [0, 2\pi)$, set $\Omega_\omega = \mathbb{C} \setminus \{re^{i\omega} : r \geq 0\}$ and show that there is a way to define $\log_\omega$ on Ω_ω so that $\exp \circ \log_\omega(z) = z$ for all $z \in \Omega_\omega$ and $\log_\omega$ is analytic on Ω_ω. What we denoted by log in (2.2.2) is one version of $\log_\pi$, and any other version differs from that one by $i2\pi n$ for some $n \in \mathbb{Z}$. The functions $\log_\omega$ are called *branches* of the logarithm function, and the one in (2.2.2) is sometimes called the *principal branch*. For each ω, the function $\log_\omega$ would have an irremovable discontinuity if one attempted to continue it across the ray $\{re^{i\omega} : r > 0\}$, and so one tries to choose the branch so that this discontinuity does the least harm. In an attempt to remove this inconvenience, Riemann introduced a beautiful geometric structure, known as a Riemann surface, that enabled him to fit all the branches together into one structure.

Chapter 3
Integration

Calculus has two components, and, thus far, we have been dealing with only one of them, namely differentiation. Differentiation is a systematic procedure for disassembling quantities at an infinitesimal level. Integration, which is the second component and is the topic here, is a systematic procedure for assembling the same sort of quantities. One of Newton's great discoveries is that these two components complement one another in a way that makes each of them more powerful.

3.1 Elements of Riemann Integration

Suppose that $f : [a, b] \longrightarrow [0, \infty)$ is a bounded function, and consider the region Ω in the plane bounded below by the portion of the horizontal axis between $(a, 0)$ and $(b, 0)$, the line segment between $(a, 0)$ and $\big(a, f(a)\big)$ on the left, the graph of f above, and the line segment between $\big(b, f(b)\big)$ and $(b, 0)$ on the right. In order to compute the area of this region, one might chop the interval $[a, b]$ into $n \geq 1$ equal parts and argue that, if f is sufficiently continuous and therefore does not vary much over small intervals, then, when n is large, the area of each slice

$$\Big\{(x, y) \in \Omega : x \in \big[a + \tfrac{m-1}{n}(b-a), a + \tfrac{m}{n}(b-a)\big] \;\&\; 0 \leq y \leq f(x)\Big\},$$

where $1 \leq m \leq n$, should be well approximated by $\frac{b-a}{n} f\big(a + \frac{m-1}{n}(b-a)\big)$, the area of the rectangle $\big[a + \frac{m-1}{n}(b-a), a + \frac{m}{n}(b-a)\big] \times \big[0, f\big(a + \frac{m-1}{n}(b-a)\big)\big]$. Continuing this line of reasoning, one would then say that the area of Ω is obtained by adding the areas of these slices and taking the limit

$$\lim_{n\to\infty} \frac{b-a}{n} \sum_{m=1}^{n} f\big(a + \tfrac{m-1}{n}(b-a)\big).$$

D.W. Stroock, *A Concise Introduction to Analysis*,
DOI 10.1007/978-3-319-24469-3_3

Of course, there are two important questions that should be asked about this procedure. In the first place, does the limit exist and, secondly, if it does, is there a compelling reason for thinking that it represents the area of Ω? Before addressing these questions, we will reformulate the preceding in a more flexible form. Say that two closed intervals are *non-overlapping* if their interiors are disjoint. Next, given a finite collection $\mathcal{C}$ of non-overlapping closed intervals $I \neq \emptyset$ whose union is $[a, b]$, a *choice function* is a map $\Xi : \mathcal{C} \longrightarrow [a, b]$ such that $\Xi(I) \in I$ for each $I \in \mathcal{C}$. Finally given $\mathcal{C}$ and Ξ, define the corresponding *Riemann sum* of a function $f : [a, b] \longrightarrow \mathbb{R}$ to be

$$\mathcal{R}(f; \mathcal{C}, \Xi) = \sum_{I \in \mathcal{C}} f\big(\Xi(I)\big)|I|, \quad \text{where } |I| \text{ is the length of } I.$$

What we want to show is that, as the *mesh size* $\|\mathcal{C}\| \equiv \max\{|I| : I \in \mathcal{C}\}$ tends to 0, for a large class of functions f these Riemann sums converge in the sense that there is a number $\int_a^b f(x)\,dx \in \mathbb{R}$, which we will call the *Riemann integral*, or simply the *integral*, of f on $[a, b]$, such that for every $\epsilon > 0$ one can find a $\delta > 0$ for which

$$\left| \mathfrak{R}(f; \mathcal{C}, \Xi) - \int_a^b f(x)\,dx \right| \leq \epsilon$$

for all $\mathcal{C}$ with $\|\mathcal{C}\| \leq \delta$ and all associated choice functions Ξ. When such a limit exists, we will say that f is *Riemann integrable* on $[a, b]$.

In order to carry out this program, it is helpful to introduce the *upper Riemann sum*

$$\mathcal{U}(f; \mathcal{C}) = \sum_{I \in \mathcal{C}} \left(\sup_I f \right) |I|, \quad \text{where } \sup_I f = \sup\{f(x) : x \in I\},$$

and the *lower Riemann sum*

$$\mathcal{L}(f; \mathcal{C}) = \sum_{I \in \mathcal{C}} \left(\inf_I f \right) |I|, \quad \text{where } \inf_I f = \inf\{f(x) : x \in I\}.$$

Lemma 3.1.1 *Assume that* $f : [a, b] \longrightarrow \mathbb{R}$ *is bounded. For every* $\mathcal{C}$ *and choice function* Ξ, $\mathcal{L}(f; \mathcal{C}) \leq \mathcal{R}(f; \mathcal{C}, \Xi) \leq \mathcal{U}(f; \mathcal{C})$. *In addition, for any pair* $\mathcal{C}$ *and* $\mathcal{C}'$, $\mathcal{L}(f; \mathcal{C}) \leq \mathcal{U}(f; \mathcal{C}')$. *Finally, for any* $\mathcal{C}$ *and* $\epsilon > 0$, *there exists a* $\delta > 0$ *such that*

$$\|\mathcal{C}'\| < \delta \implies \mathcal{L}(f; \mathcal{C}) \leq \mathcal{L}(f; \mathcal{C}') + \epsilon \text{ and } \mathcal{U}(f; \mathcal{C}') \leq \mathcal{U}(f; \mathcal{C}) + \epsilon,$$

and therefore

$$\lim_{\|\mathcal{C}\| \to 0} \mathcal{L}(f; \mathcal{C}) = \sup_{\mathcal{C}} \mathcal{L}(f; \mathcal{C}) \leq \inf_{\mathcal{C}} \mathcal{U}(f; \mathcal{C}) = \lim_{\|\mathcal{C}\| \to 0} \mathcal{U}(f; \mathcal{C}).$$

Proof The first assertion is obvious. To prove the second, begin by observing that there is nothing to do if $\mathcal{C}' = \mathcal{C}$. Next, suppose that every $I \in \mathcal{C}$ is contained in some $I' \in \mathcal{C}'$, in which case $\sup_I f \le \sup_{I'} f$. Therefore (cf. Lemma 5.1.1 for a detailed proof), since each $I' \in \mathcal{C}'$ is the union of the I's in $\mathcal{C}$ which it contains,

$$\begin{aligned}\mathcal{U}(f;\mathcal{C}') &= \sum_{I'\in\mathcal{C}'}\left(\sum_{\substack{I\in\mathcal{C}\\ I\subseteq I'}}\left(\sup_{I'} f\right)|I|\right)\\ &\ge \sum_{I'\in\mathcal{C}'}\left(\sum_{\substack{I\in\mathcal{C}\\ I\subseteq I'}}\left(\sup_{I} f\right)|I|\right) = \sum_{I\in\mathcal{C}}\left(\sup_{I} f\right)|I| = \mathcal{U}(f;\mathcal{C}),\end{aligned}$$

and, similarly, $\mathcal{L}(f;\mathcal{C}') \le \mathcal{L}(f;\mathcal{C})$. Now suppose that $\mathcal{C}$ and $\mathcal{C}'$ are given, and set

$$\mathcal{C}'' \equiv \{I \cap I' : I \in \mathcal{C},\ I' \in \mathcal{C}' \ \&\ I \cap I' \ne \emptyset\}.$$

Then for each I'' there exist $I \in \mathcal{C}$ and $I' \in \mathcal{C}'$ such that $I'' \subseteq I$ and $I'' \subseteq I'$, and therefore

$$\mathcal{L}(f;\mathcal{C}) \le \mathcal{L}(f;\mathcal{C}'') \le \mathcal{U}(f;\mathcal{C}'') \le \mathcal{U}(f,\mathcal{C}').$$

Finally, let $\mathcal{C}$ and $\epsilon > 0$ be given, and choose $a = c_0 \le c_1 \le \cdots \le c_K = b$ so that $\mathcal{C} = \{[c_{k-1}, c_k] : 1 \le k \le K\}$. Given $\mathcal{C}'$, let $\mathcal{D}$ be the set of $I' \in \mathcal{C}'$ with the property that $c_k \in \operatorname{int}(I')$ for at least one $1 \le k < K$, and observe that, since the intervals are non-overlapping, $\mathcal{D}$ can contain at most $K-1$ elements. Then each $I' \in \mathcal{C}' \setminus \mathcal{D}$ is contained in some $I \in \mathcal{C}$ and therefore $\sup_{I'} f \le \sup_I f$. Hence

$$\begin{aligned}\mathcal{U}(f;\mathcal{C}') &= \sum_{I'\in\mathcal{D}}\left(\sup_{I'} f\right)|I'| + \sum_{I\in\mathcal{C}}\sum_{\substack{I'\in\mathcal{C}'\setminus\mathcal{D}\\ I'\subseteq I}}\left(\sup_{I'} f\right)|I'|\\ &\le \left(\sup_{[a,b]}|f|\right)(K-1)\|\mathcal{C}'\| + \sum_{I\in\mathcal{C}}\sup_I f\left(\sum_{\substack{I'\in\mathcal{C}'\\ I'\subseteq I}}|I'|\right)\\ &\le \left(\sup_{[a,b]}|f|\right)(K-1)\|\mathcal{C}'\| + \mathcal{U}(f;\mathcal{C}).\end{aligned}$$

Therefore, if $\delta > 0$ is chosen so that $\left(\sup_{[a,b]}|f|\right)(K-1)\delta < \epsilon$, then $\mathcal{U}(f;\mathcal{C}') \le \mathcal{U}(f;\mathcal{C}) + \epsilon$ whenever $\|\mathcal{C}'\| \le \delta$. Applying this to $-f$ and noting that $\mathcal{L}(f;\mathcal{C}'') = -\mathcal{U}(-f;\mathcal{C}'')$ for any $\mathcal{C}''$, one also has that $\mathcal{L}(f;\mathcal{C}) \le \mathcal{L}(f;\mathcal{C}') + \epsilon$ if $\|\mathcal{C}'\| \le \delta$. $\square$

Theorem 3.1.2 *If $f : [a, b] \longrightarrow \mathbb{R}$ is a bounded function, then it is Riemann integrable if and only if for each $\epsilon > 0$ there is a $\mathcal{C}$ such that*

$$\sum_{\substack{I \in \mathcal{C} \\ \sup_I f - \inf_I f \geq \epsilon}} |I| < \epsilon.$$

In particular, a bounded f will be Riemann integrable if it is continuous at all but a finite number of points. In addition, if $f : [a, b] \longrightarrow [c, d]$ is Riemann integrable and $\varphi : [c, d] \longrightarrow \mathbb{R}$ is continuous, then $\varphi \circ f$ is Riemann integrable on $[a, b]$.

Proof First assume that f is Riemann integrable. Given $\epsilon > 0$, choose $\mathcal{C}$ so that

$$\left| \mathcal{R}(f; \mathcal{C}, \Xi) - \int_a^b f(x)\, dx \right| < \frac{\epsilon^2}{6}$$

for all choice functions Ξ. Next choose choice functions $\Xi^\pm$ so that

$$f\big(\Xi^+(I)\big) + \frac{\epsilon^2}{3(b-a)} \geq \sup_I f \quad \text{and} \quad f\big(\Xi^-(I)\big) - \frac{\epsilon^2}{3(b-a)} \leq \inf_I f$$

for each $I \in \mathcal{C}$. Then

$$\mathcal{U}(f; \mathcal{C}) \leq \mathcal{R}(f; \mathcal{C}, \Xi^+) + \frac{\epsilon^2}{3} \leq \mathcal{R}(f; \mathcal{C}, \Xi^-) + \frac{2\epsilon^2}{3} \leq \mathcal{L}(f; \mathcal{C}) + \epsilon^2,$$

and so

$$\epsilon^2 \geq \mathcal{U}(f; \mathcal{C}) - \mathcal{L}(f; \mathcal{C}) = \sum_{I \in \mathcal{C}} \left(\sup_I f - \inf_I f \right) |I| \geq \epsilon \sum_{\substack{I \in \mathcal{C} \\ \sup_I f - \inf_I f \geq \epsilon}} |I|.$$

Next assume that for each $\epsilon > 0$ there is a $\mathcal{C}$ such that

$$\sum_{\substack{I \in \mathcal{C} \\ \sup_I f - \inf_I f \geq \epsilon}} |I| < \epsilon.$$

Given $\epsilon > 0$, set $\epsilon' = \epsilon \left(4 \sup_{[a,b]} |f| + 2(b-a)\right)^{-1}$, and choose $\mathcal{C}_\epsilon$ so that

$$\sum_{\substack{I \in \mathcal{C}_\epsilon \\ \sup_I f - \inf_I f \geq \epsilon'}} |I| < \epsilon'$$

and therefore

$$\mathcal{U}(f;\mathcal{C}_\epsilon) - \mathcal{L}(f;\mathcal{C}_\epsilon) \le \left(2\sup_{[a,b]}|f|\right) \sum_{\substack{I\in\mathcal{C}_\epsilon \\ \sup_I f - \inf_I f \ge \epsilon'}} |I| + \epsilon'(b-a) = \frac{\epsilon}{2}.$$

Now, using Lemma 3.1.1, choose $\delta_\epsilon > 0$ so that $\mathcal{U}(f;\mathcal{C}) \le \mathcal{U}(f;\mathcal{C}_\epsilon) + \frac{\epsilon}{2}$ and $\mathcal{L}(f;\mathcal{C}) \ge \mathcal{L}(f;\mathcal{C}_\epsilon) - \frac{\epsilon}{2}$ when $\|\mathcal{C}\| < \delta_\epsilon$. Then

$$\|\mathcal{C}\| < \delta_\epsilon \implies \mathcal{U}(f;\mathcal{C}) \le \mathcal{L}(f;\mathcal{C}) + \epsilon,$$

and so, in conjunction with Lemma 3.1.1, we know that

$$\lim_{\|\mathcal{C}\|\to 0} \mathcal{U}(f;\mathcal{C}) = M = \lim_{\|\mathcal{C}\|\to 0} \mathcal{L}(f;\mathcal{C}) \quad \text{where } M = \sup_{\mathcal{C}} \mathcal{L}(f;\mathcal{C}).$$

Since, for any $\mathcal{C}$ and associated Ξ, $\mathcal{L}(f;\mathcal{C}) \le \mathfrak{R}(f;\mathcal{C},\Xi) \le \mathcal{U}(f;\mathcal{C})$, it follows that f is Riemann integral and that M is its integral.

Turning to the next assertion, first suppose that f is continuous on $[a,b]$. Then, because it is uniformly continuous there, for each $\epsilon > 0$ there exists a $\delta_\epsilon > 0$ such that $|f(y) - f(x)| < \epsilon$ whenever $x,\ y \in [a,b]$ and $|y-x| \le \delta_\epsilon$. Hence, if $\|\mathcal{C}\| < \delta_\epsilon$, then $\sup_I f - \inf_I f < \epsilon$ for all $I \in \mathcal{C}$, and so

$$\sum_{\substack{I\in\mathcal{C} \\ \sup_I f - \inf_I f \ge \epsilon}} |I| = 0.$$

Now suppose that f is continuous except at the points $a \le c_0 < \cdots < c_K \le b$. For each $r > 0$, f is uniformly continuous on $F_r \equiv [a,b] \setminus \bigcup_{k=0}^K (c_k - r, c_k + r)$. Given $\epsilon > 0$, choose $0 < r < \min\{c_k - c_{k-1} : 1 \le k \le K\}$ so that $2r(K+1) < \epsilon$, and then choose $\delta > 0$ so that $|f(y) - f(x)| < \epsilon$ for $x,\ y \in F_r$ with $|y-x| \le \delta$. Then one can easily construct a $\mathcal{C}$ such that $I_k \equiv [c_k - r, c_k + r] \cap [a,b] \in \mathcal{C}$ for each $0 \le k \le K$ and all the other I's in $\mathcal{C}$ have length less than δ, and for such a $\mathcal{C}$

$$\sum_{\substack{I\in\mathcal{C} \\ \sup_I f - \inf_I f \ge \epsilon}} |I| \le \sum_{k=0}^K |I_k| \le 2r(K+1) < \epsilon.$$

Finally, to prove the last assertion, let $\epsilon > 0$ be given and choose $0 < \epsilon' \le \epsilon$ so that $|\varphi(\eta) - \varphi(\xi)| < \epsilon$ if $\xi,\ \eta \in [c,d]$ with $|\eta - \xi| < \epsilon'$. Next choose $\mathcal{C}$ so that

$$\sum_{\substack{I\in\mathcal{C} \\ \sup_I f - \inf_I f \ge \epsilon'}} |I| < \epsilon',$$

and conclude that

$$\sum_{\substack{I\in\mathcal{C}\\ \sup_I \varphi\circ f-\inf_I \varphi\circ f\geq\epsilon}} |I| \leq \sum_{\substack{I\in\mathcal{C}\\ \sup_I f-\inf_I f\geq\epsilon'}} |I| \leq \epsilon' \leq \epsilon.$$

□

The fact that a bounded function is Riemann integrable if it is continuous at all but a finite number of points is important. For example, if f is a bounded, continuous function on (a, b), then its integral on $[a, b]$ can be unambiguously defined by extending f to $[a, b]$ in any convenient manner (e.g. taking $f(a) = f(b) = 0$), and then taking its integral to be the integral of the extension. The result will be the same no matter how the extension is made.

When applied to a non-negative, Riemann integrable function f on $[a, b]$, Theorem 3.1.2 should be convincing evidence that the procedure we suggested for computing the area of the region Ω gives the correct result. Indeed, given any $\mathcal{C}$, $\mathcal{U}(f;\mathcal{C})$ dominates the area of Ω and $\mathcal{L}(f;\mathcal{C})$ is dominated by the area of Ω. Hence, since by taking $\|\mathcal{C}\|$ small we can close the gap between $\mathcal{U}(f;\mathcal{C})$ and $\mathcal{L}(f;\mathcal{C})$, there can be little doubt that $\int_a^b f(x)\,dx$ is the area of Ω. More generally, when f takes both signs, one can interpret $\int_a^b f(x)\,dx$ as the difference between the area above the horizontal axis and the area below.

The following corollary deals with several important properties of Riemann integrals. In its statement and elsewhere, if $f : S \longrightarrow \mathbb{C}$,

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

is the *uniform norm* of f on S. Observe that

$$\|fg\|_S \leq \|f\|_S\|g\|_S \text{ and } \|\alpha f + \beta g\|_S \leq |\alpha|\|f\|_S + |\beta|\|g\|_S$$

for all $\mathbb{C}$-valued functions f and g on S and $\alpha, \beta \in \mathbb{C}$.

Corollary 3.1.3 *If $f : [a, b] \longrightarrow \mathbb{R}$ is a bounded, Riemann integrable function and $a < c < b$, then f is Riemann integrable on both $[a, c]$ and $[c, b]$, and*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b g(x)\,dx \quad \textit{for all } c \in (a, b). \tag{3.1.1}$$

Further, if $\lambda > 0$ and f is a bounded, Riemann integrable function on $[\lambda a, \lambda b]$, then $x \in [a, b] \longmapsto f(\lambda x) \in \mathbb{R}$ is Riemann integrable and

$$\int_{\lambda a}^{\lambda b} f(y)\,dy = \lambda \int_a^b f(\lambda x)\,dx. \tag{3.1.2}$$

Next suppose that f and g are bounded, Riemann integrable functions on $[a, b]$. Then,

$$f \le g \implies \int_a^b f(x)\,dx \le \int_a^b g(x)\,dx, \tag{3.1.3}$$

and so, if f is a bounded, Riemann integrable function on $[a, b]$, then

$$\left|\int_a^b f(x)\,dx\right| \le \left|\int_a^b |f(x)|\,dx\right| \le \|f\|_S|b-a|. \tag{3.1.4}$$

In addition, fg is Riemann integrable on $[a, b]$, and, for all $\alpha,\ \beta \in \mathbb{R}$, $\alpha f + \beta g$ is also Riemann integrable and

$$\int_a^b \big(\alpha f(x) + \beta g(x)\big)\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx. \tag{3.1.5}$$

Proof To prove the first assertion, for a given $\epsilon > 0$ choose a non-overlapping cover $\mathcal{C}$ of $[a, b]$ so that

$$\sum_{\substack{I\in\mathcal{C}\\ \sup_I f - \inf_I f \ge \epsilon}} |I| < \epsilon,$$

and set $\mathcal{C}' = \{I \cap [a, c] : I \in \mathcal{C}\}$. Then, since

$$\sup\{f(y) - f(x) : x,\ y \in I \cap [a, c]\} \le \sup\{f(y) - f(x) : x,\ y \in I\}$$

and $|I \cap [a, c]| \le |I|$,

$$\sum_{\substack{I'\in\mathcal{C}'\\ \sup_{I'} f - \inf_{I'} f \ge \epsilon}} |I'| \le \sum_{\substack{I\in\mathcal{C}\\ \sup_I f - \inf_I f \ge \epsilon}} |I| < \epsilon.$$

Thus, f is Riemann integrable on $[a, c]$. The proof that f is also Riemann integrable on $[c, b]$ is the same. As for (3.1.1), choose $\{\mathcal{C}_n : n \ge 1\}$ and $\{\mathcal{C}'_n : n \ge 1\}$ and associated choice functions $\{\Xi_n : n \ge 1\}$ and $\{\Xi'_n : n \ge 1\}$ for $[a, c]$ and $[c, b]$ so that $\|\mathcal{C}_n\| \vee \|\mathcal{C}'_n\| \le \frac{1}{n}$, and set $\mathcal{C}''_n = \mathcal{C}_n \cup \mathcal{C}'_n$ and $\Xi''_n(I) = \Xi_n(I)$ if $I \in \mathcal{C}_n$ and $\Xi''_n(I) = \Xi'_n(I)$ if $I \in \mathcal{C}'_n \setminus \mathcal{C}_n$. Then

$$\begin{aligned}\int_a^b f(x)\,dx &= \lim_{n\to\infty} \mathfrak{R}(f; \mathcal{C}''_n, \Xi''_n) = \lim_{n\to\infty} \mathfrak{R}(f; \mathcal{C}_n, \Xi_n) + \lim_{n\to\infty} \mathfrak{R}(f; \mathcal{C}'_n, \Xi'_n)\\ &= \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.\end{aligned}$$

Turning to the second assertion, set $f_\lambda(x) = f(\lambda x)$ for $x \in [a, b]$ and, given a cover $\mathcal{C}$ and a choice function Ξ, take $I_\lambda = \{\lambda x : x \in I\}$, and define

$\mathcal{C}_\lambda = \{I_\lambda : I \in \mathcal{C}\}$ and $\Xi_\lambda(I_\lambda) = \lambda\Xi(I)$ for $I \in \mathcal{C}$. Then $\mathcal{C}_\lambda$ is a non-overlapping cover for $[\lambda a, \lambda b]$, Ξ_λ is an associated choice function, $\|\mathcal{C}_\lambda\| = \lambda\|\mathcal{C}\|$, and

$$\mathcal{R}(f_\lambda; \mathcal{C}, \Xi) = \sum_{I\in\mathcal{C}} f\big(\lambda\Xi(I)\big)|I| = \lambda^{-1}\mathcal{R}(f; \mathcal{C}_\lambda, \Xi_\lambda) \longrightarrow \lambda^{-1}\int_{\lambda a}^{\lambda b} f(y)\,dy$$

as $\|\mathcal{C}\| \to 0$.

Next suppose that f and g are bounded, Riemann integrable functions on $[a, b]$. Obviously, for all $\mathcal{C}$ and Ξ, $\mathcal{R}(f; \mathcal{C}, \Xi) \le \mathcal{R}(g; \mathcal{C}, \Xi)$ if $f \le g$ and

$$\mathcal{R}(\alpha f + \beta g; \mathcal{C}, \Xi) = \alpha\mathcal{R}(f; \mathcal{C}, \Xi) + \beta\mathcal{R}(g; \mathcal{C}, \Xi)$$

for all $\alpha, \beta \in \mathbb{R}$. Starting from these and passing to the limit as $\|\mathcal{C}\| \to 0$, one arrives at the conclusions in (3.1.3) and (3.1.5). Furthermore, (3.1.4) follows from (3.1.3), since, by the last part of Theorem 3.1.2, $|f|$ is Riemann integrable and $\pm f \le |f| \le \|f\|_{[a,b]}$. Finally, to see that fg is Riemann integrable, note that, by the preceding combined with the last part of Theorem 3.1.2, $(f+g)^2$ and $(f-g)^2$ are both Riemann integrable and therefore so is $fg = \frac{1}{4}\big((f+g)^2 - (f-g)^2\big)$. $\square$

There is a useful notation convention connected to (3.1.1). Namely, if $a < b$, then one defines

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx.$$

With this convention, if $\{a_k : 0 \le k \le \ell\} \subseteq \mathbb{R}$ and f is function that is Riemann integrable on $[a_0, a_\ell]$ and $[a_k, a_{k+1}]$ for each $0 \le k < \ell$, one can use (3.1.1) to check that

$$\int_{a_0}^{a_\ell} f(x)\,dx = \sum_{k=0}^{\ell-1}\int_{a_k}^{a_{k+1}} f(x)\,dx. \tag{3.1.6}$$

Sometimes one wants to integrate functions that are complex-valued. Just as in the real-valued case, one says that a bounded function $f : [a, b] \longrightarrow \mathbb{C}$ is *Riemann integrable* and has *integral* $\int_a^b f(x)\,dx$ on $[a, b]$ if the Riemann sums

$$\mathcal{R}(f; \mathcal{C}, \Xi) = \sum_{I\in\mathcal{C}} f\big(\Xi(I)\big)|I|$$

converge to $\int_a^b f(x)\,dx$ as $\|\mathcal{C}\| \to 0$. Writing $f = u + iv$, where u and v are real-valued, one can easily check that f is Riemann integrable if and only if both u and v are, in which case

$$\int_a^b f(x)\,dx = \int_a^b u(x)\,dx + i\int_a^b v(x)\,dx.$$

From this one sees that, except for (3.1.3), the obvious analogs of the assertions in Corollary 3.1.3 continue to hold for complex-valued functions. Of course, one can no longer use (3.1.3) to prove the first inequality in (3.1.4). Instead, one can use the triangle inequality to show that $|\mathfrak{R}(f;\mathcal{C},\Xi)| \le \mathfrak{R}(|f|;\mathcal{C},\Xi)$ and then take the limit as $\|\mathcal{C}\| \to 0$.

There are two closely related extensions of the Riemann integral. In the first place, one often has to deal with an interval $(a,b]$ on which there is a function f that is unbounded but is bounded and Riemann integrable on $[\alpha,b]$ for each $\alpha \in (a,b)$. Even though f is unbounded, it may be that the limit $\lim_{\alpha\searrow a}\int_\alpha^b f(x)\,dx$ exists in $\mathbb{C}$, in which case one uses $\int_{(a,b]} f(x)\,dx$ to denote the limit. Similarly, if f is bounded and Riemann integrable on $[a,\beta]$ for each $\beta \in (a,b)$ and $\lim_{\beta\nearrow b}\int_a^\beta f(x)\,dx$ exists or if f is bounded and Riemann integrable on $[\alpha,\beta]$ for all $a < \alpha < \beta < b$ and $\lim_{\substack{\alpha\searrow a\\ \beta\nearrow b}}\int_\alpha^\beta f(x)\,dx$ exists in $\mathbb{C}$, then one takes $\int_{[a,b)} f(x)\,dx$ or $\int_{(a,b)} f(x)\,dx$ to be the corresponding limit. The other extension is to situations when one or both of the endpoints are infinite. In this case one is dealing with a function f which is bounded and Riemann integrable on bounded, closed intervals of $(-\infty,b]$, $[a,\infty)$, or $(-\infty,\infty)$, and one takes

$$\int_{(-\infty,b]} f(x)\,dx, \quad \int_{[a,\infty)} f(x)\,dx, \quad \text{or} \int_{(-\infty,\infty)} f(x)\,dx$$

to be

$$\lim_{a\searrow-\infty}\int_a^b f(x)\,dx, \ \lim_{b\nearrow\infty}\int_a^b f(x)\,dx, \ \text{or} \ \lim_{\substack{b\nearrow\infty\\ a\searrow-\infty}}\int_a^b f(x)\,dx$$

if the corresponding limit exists. Notice that in any of these situations, if f is non-negative then the corresponding limits exist in $[0,\infty)$ if and only if the quantities of which one is taking the limit stay bounded. More generally, if f is a $\mathbb{C}$-valued function on an interval J and if f is Riemann integrable on each bounded, closed interval $I \subseteq J$, then $\int_J f(x)\,dx$ will exist if $\sup_I \int_I |f(x)|\,dx < \infty$, in which is case f is said to be *absolutely Riemann integrable* on J. To check this, suppose that $J = (a,b]$. Then, for $a < \alpha_1 < \alpha_2 < b$,

$$\left|\int_{\alpha_2}^b f(x)\,dx - \int_{\alpha_1}^b f(x)\,dx\right| = \left|\int_{\alpha_1}^{\alpha_2} f(x)\,dx\right| \le \int_{\alpha_1}^{\alpha_2} |f(x)|\,dx,$$

and so the existence of the limit $\lim_{\alpha\searrow a}\int_\alpha^b f(x)\,dx \in \mathbb{C}$ follows from the existence of $\lim_{\alpha\searrow a}\int_\alpha^b |f(x)|\,dx \in [0,\infty)$. When J is $[a,b)$, (a,b), $[a,\infty)$, $(-\infty,b]$, or $(-\infty,\infty)$, the argument is essentially the same.

The following theorem shows that integrals are continuous with respect to uniform convergence of their integrands.

Theorem 3.1.4 *If* $\{f_n : n \geq 1\}$ *is a sequence of bounded, Riemann integrable* $\mathbb{C}$-*valued functions on* $[a, b]$ *that converge uniformly to the function* $f : [a, b] \longrightarrow \mathbb{C}$, *then* f *is Riemann integrable and*

$$\lim_{n\to\infty} \int_a^b f_n(x)\, dx = \int_a^b f(x)\, dx.$$

Proof Observe that

$$|\mathcal{R}(f_n; \mathcal{C}, \Xi) - \mathcal{R}(f; \mathcal{C}, \Xi)| \leq \mathcal{R}\big(|f_n - f|; \mathcal{C}, \Xi\big) \leq (b-a)\|f_n - f\|_{[a,b]},$$

and conclude from this first that $\left\{\int_a^b f_n(x)\, dx : n \geq 1\right\}$ satisfies Cauchy's convergence criterion and second that, for each n,

$$\varlimsup_{\|\mathcal{C}\|\to 0} \left|\mathcal{R}(f; \mathcal{C}, \Xi) - \int_a^b f_n(x)\, dx\right| \leq (b-a)\|f_n - f\|_{[a,b]}.$$

Hence, $\int_a^b f(x)\, dx \equiv \lim_{n\to\infty} \int_a^b f_n(x)\, dx$ exists, and $\mathcal{R}(f; \mathcal{C}, \Xi) \longrightarrow \int_a^b f(x)\, dx$ as $\|\mathcal{C}\| \to 0$. □

3.2 The Fundamental Theorem of Calculus

In the preceding section we developed a lot of theory for integrals but did not address the problem of actually evaluating them. To see that the theory we have developed thus far does little to make computations easy, consider the problem of computing $\int_a^b x^k\, dx$ for $k \in \mathbb{N}$. When $k = 0$, it is obvious that every Riemann sum will be equal to $(b - a)$, and therefore $\int_a^b x^0\, dx = b - a$. To handle $k \geq 1$, first note that $\int_a^b x^k\, dx = \int_0^b x^k\, dx - \int_0^a x^k\, dx$ and, for any $c \in \mathbb{R}$

$$\int_0^{-c} x^k\, dx = (-1)^{k+1} \int_0^c x^k\, dx.$$

Hence, it suffices to compute $\int_0^c x^k\, dx$ for $c > 0$. Further, by the scaling property in (3.1.2),

$$\int_0^c x^k\, dx = c^{k+1} \int_0^1 x^k\, dx.$$

Thus, everything comes down to computing $\int_0^1 x^k\, dx$. To this end, we look at Riemann sums of the form

$$\frac{1}{n}\sum_{m=1}^{n}\left(\frac{m}{n}\right)^k = \frac{S_n^{(k)}}{n^{k+1}} \quad \text{where } S_n^{(k)} \equiv \sum_{m=1}^{n} m^k.$$

When n is even, one can see that $S_n^{(1)} = \frac{n(n+1)}{2}$ by adding the 1 to n, 2 to $n-1$, etc. When n is odd, one gets the same conclusion by adding n to $S_{n-1}^{(1)}$. Hence, we have shown that

$$\int_0^1 x\,dx = \lim_{n\to\infty}\frac{n(n+1)}{2n^2} = \frac{1}{2},$$

which is what one would hope is the area of the right triangle with vertices $(0, 0)$, $(0, 1)$, and $(1, 1)$. When $k \geq 2$ one can proceed as follows. Write the difference $(n+1)^{k+1} - 1$ as the telescoping sum $\sum_{m=1}^{n}\big((m+1)^{k+1} - m^{k+1}\big)$. Next, expand $(m+1)^{k+1}$ using the binomial formula, and, after changing the order of summation, arrive at

$$(n+1)^{k+1} - 1 = \sum_{j=0}^{k}\binom{k+1}{j} S_n^{(j)}.$$

Starting from this and using induction on k, one sees that $\lim_{n\to\infty}\frac{S_n^{(k)}}{n^{k+1}} = \frac{1}{k+1}$. Hence, we now know that $\int_0^1 x^k\,dx = \frac{1}{k+1}$. Combining this with our earlier comments, we have

$$\int_a^b x^k\,dx = \frac{b^{k+1} - a^{k+1}}{k+1} \quad \text{for all } k \in \mathbb{N} \text{ and } a < b. \tag{3.2.1}$$

There is something that should be noticed about the result (3.2.1). Namely, if one looks at $\int_a^b x^k\,dx$ as a function of the upper limit b, then b^k is its derivative. That is, as a function of its upper limit, the derivative of this integral is the integrand. That this is true in general is one of Newton's great discoveries.[1]

Theorem 3.2.1 (Fundamental Theorem of Calculus) *Let $f : [a, b] \longrightarrow \mathbb{C}$ be a continuous function, and set $F(x) = \int_a^x f(t)\,dt$ for $x \in [a, b]$. Then F is continuous on $[a, b]$, continuously differentiable on (a, b), and $F' = f$ there. Conversely, if $F : [a, b] \longrightarrow \mathbb{C}$ is a continuous function that is continuously differentiable on (a, b), then*

$$F' = f \text{ on } (a, b) \implies F(b) - F(a) = \int_a^b f(x)\,dx.$$

[1] Although it was Newton who made this result famous, it had antecedents in the work of James Gregory and Newton's teacher Isaac Barrow. Mathematicians are not always reliable historians, and their attributions should be taken with a grain of salt.

Proof Without loss in generality, we will assume that f is $\mathbb{R}$-valued.

Let F be as in the first assertion. Then, by (3.1.1), for $x,\ y \in [a, b]$,

$$F(y) - F(x) = \int_x^y f(t)\,dt = f(x)(y-x) + \int_x^y \big(f(t) - f(x)\big)\,dt.$$

Given $\epsilon > 0$, choose $\delta > 0$ so that $|f(t) - f(x)| < \epsilon$ for $t \in [a, b]$ with $|t - x| < \delta$. Then, by (3.1.4),

$$\left|\int_x^y \big(f(t) - f(x)\big)\,dt\right| < \epsilon|y - x| \quad \text{if } |y - x| < \delta,$$

and so

$$\left|\frac{F(y) - F(x)}{y - x} - f(x)\right| < \epsilon \quad \text{for } x,\ y \in [a, b] \text{ with } 0 < |y - x| < \delta.$$

This proves that F is continuous on $[a, b]$, differentiable on (a, b), and $F' = f$ there.

If f and F are as in the second assertion, set $\Delta(x) = F(x) - \int_a^x f(t)\,dt$. Then, by the preceding, Δ is continuous of $[a, b]$, differentiable on (a, b), and $\Delta' = 0$ on (a, b). Hence, by (1.8.1), $\Delta(b) = \Delta(a)$, and so, since $\Delta(a) = F(a)$, the asserted result follows. □

It is hard to overstate the importance of Theorem 3.2.1. The hands-on method we used to integrate x^k is unable to handle more complicated functions. Instead, given a function f, one looks for a function F such that $F' = f$ and then applies Theorem 3.2.1 to do the calculation. Such a function F is called an *indefinite integral* of f. By Theorem 1.8.1, since the derivative of the difference between any two of its indefinite integrals is 0, two indefinite integrals of a function can differ by at most an additive constant. Once one has F, it is customary to write

$$\int_a^b f(x)\,dx = F(x)\Big|_{x=a}^b \equiv F(b) - F(a).$$

Here are a couple of corollaries of Theorem 3.2.1.

Corollary 3.2.2 *Suppose that f and g are continuous, $\mathbb{C}$-valued functions on $[a, b]$ which are continuously differentiable on (a, b), and assume that their derivatives are bounded. Then*

$$\int_a^b f'(x)g(x)\,dx = f(x)g(x)\Big|_{x=a}^b - \int_a^b f(x)g'(x)\,dx$$

where $f(x)g(x)\Big|_{x=a}^b \equiv f(b)g(b) - f(a)g(a)$.

Proof By the product rule, $(fg)' = f'g + gf'$, and so, by Theorem 3.2.1 and (3.1.5),

$$f(\beta)g(\beta) - f(\alpha)g(\alpha) = \int_\alpha^\beta f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx$$

for all $a < \alpha < \beta < b$. Now let $\alpha \searrow a$ and $\beta \nearrow b$. □

The equation in Corollary 3.2.2 is known as the *integration by parts formula*, and it is among the most useful tools available for computing integrals. For instance, it can be used to give another derivation of Taylor's theorem, this time with the remainder term expressed as an integral. To be precise, let $f : (a, b) \longrightarrow \mathbb{C}$ be an $(n+1)$ times continuously differentiable function. Then, for $x,\ y \in (a, b)$,

$$\begin{aligned} f(y) = &\sum_{m=0}^{n} \frac{(y-x)^m}{m!} f^{(m)}(x) \\ &+ \frac{(y-x)^{n+1}}{n!} \int_0^1 (1-t)^n f^{(n+1)}\big((1-t)x + ty\big)\,dt. \end{aligned} \tag{3.2.2}$$

To check this, set $u(t) = f\big((1-t)x + ty\big)$. Then

$$u^{(m)}(t) = (y-x)^m f^{(m)}\big((1-t)x + ty\big),$$

and, by Theorem 3.2.1, $u(1) - u(0) = \int_0^1 u'(t)\,dt$, which is (3.2.2) for $n = 1$. Next, assume that

$$(*) \qquad u(1) = \sum_{m=0}^{n-1} \frac{u^{(m)}(0)}{m!} + \frac{1}{(n-1)!} \int_0^1 (1-t)^{n-1} u^{(n)}(t)\,dt$$

for some $n \geq 1$, and use integration by parts to see that

$$\int_0^1 (1-t)^{n-1} u^{(n)}(t)\,dt = -\frac{(1-t)^n u^{(n)}(t)}{n}\Big|_{t=0}^{1} + \frac{1}{n}\int_0^1 (1-t)^n u^{(n+1)}(t)\,dt.$$

Hence, by induction, (3.2.2) holds for all $n \geq 1$.

A second application of integration by parts is to the derivation of *Wallis's formula*:

$$\frac{\pi}{2} = \lim_{n\to\infty} \prod_{m=1}^{n} \frac{2m}{2m-1}\frac{2m}{2m+1} = \lim_{n\to\infty} \frac{4^n (n!)^2}{(2n+1)\left(\prod_{m=1}^{n}(2m-1)\right)^2}. \tag{3.2.3}$$

To prove his formula, we begin by using integration by parts to see that

$$\begin{aligned}\int_0^{\frac{\pi}{2}} \cos^m t\, dt &= \int_0^{\frac{\pi}{2}} \sin' t \cos^{m-1} t\, dt = (m-1)\int_0^{\frac{\pi}{2}} \sin^2 t \cos^{m-2} t\, dt \\ &= (m-1)\int_0^{\frac{\pi}{2}} \cos^{m-2} t\, dt - (m-1)\int_0^{\frac{\pi}{2}} \cos^m t\, dt,\end{aligned}$$

and therefore that

$$\int_0^{\frac{\pi}{2}} \cos^m t\, dt = \frac{m-1}{m}\int_0^{\frac{\pi}{2}} \cos^{m-2} t\, dt \quad \text{for } m \geq 2.$$

Since $\int_0^{\frac{\pi}{2}} \cos t\, dt = 1$, this proves that

$$\int_0^{\frac{\pi}{2}} \cos^{2n+1} t\, dt = \prod_{m=1}^{n} \frac{2m}{2m+1} \quad \text{for } n \geq 1.$$

At the same time, it shows that $\int_0^{\frac{\pi}{2}} \cos^2 t\, dt = \frac{\pi}{4}$ and therefore that

$$\int_0^{\frac{\pi}{2}} \cos^{2n} t\, dt = \frac{\pi}{2}\prod_{m=1}^{n} \frac{2m-1}{2m} \quad \text{for } n \geq 1.$$

Thus

$$\frac{\int_0^{\frac{\pi}{2}} \cos^{2n+1} t\, dt}{\int_0^{\frac{\pi}{2}} \cos^{2n} t\, dt} = \frac{2}{\pi}\frac{4^n (n!)^2}{(2n+1)\left(\prod_{m=1}^{n}(2m-1)\right)^2}.$$

Finally, since

$$1 \geq \frac{\int_0^{\frac{\pi}{2}} \cos^{2n+1} t\, dt}{\int_0^{\frac{\pi}{2}} \cos^{2n} t\, dt} = \frac{2n}{2n+1}\frac{\int_0^{\frac{\pi}{2}} \cos^{2n-1} t\, dt}{\int_0^{\frac{\pi}{2}} \cos^{2n} t\, dt} \geq \frac{2n}{2n+1},$$

(3.2.3) follows.

Aside from being a curiosity, as Stirling showed, Wallis's formula allows one to compute the constant e^{Δ} in (1.8.7). To understand what he did, observe that $\prod_{m=1}^{n}(2m-1) = \frac{(2n)!}{2^n n!}$ and therefore, by (1.8.7), that

$$\begin{aligned}\frac{4^n (n!)^2}{(2n+1)\left(\prod_{m=1}^{n+1}(2m-1)\right)^2} &= \frac{1}{2n+1}\left(\frac{4^n (n!)^2}{(2n)!}\right)^2 \\ &\sim \frac{1}{2n+1}\left(\frac{4^n e^{\Delta} n \left(\frac{n}{e}\right)^{2n}}{\sqrt{2n}\left(\frac{2n}{e}\right)^{2n}}\right)^2 = \frac{e^{2\Delta} n}{4n+2}.\end{aligned}$$

Hence, after letting $n \to \infty$ and applying (3.2.3), one sees that $e^{2\Delta} = 2\pi$ and therefore that (1.8.7) can be replaced by

$$\sqrt{2\pi} e^{-\frac{1}{n}} \le \frac{n! e^n}{n^{n+\frac{1}{2}}} \le \sqrt{2\pi} e^{\frac{1}{n}}, \tag{3.2.4}$$

or, more imprecisely, $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.

Here is another powerful tool for computing integrals.

Corollary 3.2.3 *Let $\varphi : [a, b] \longrightarrow [c, d]$ be a continuous function, and assume that φ is continuously differentiable on (a, b) and that its derivative is bounded. If $f : [c, d] \longrightarrow \mathbb{C}$ is a continuous function, then*

$$\int_a^b \big(f \circ \varphi(t)\big)\varphi'(t)\, dt = \int_{\varphi(a)}^{\varphi(b)} f(x)\, dx.$$

In particular, if $\varphi' > 0$ on (a, b) and $\varphi(a) \le c < d \le \varphi(b)$, then for any continuous $f : [c, d] \longrightarrow \mathbb{C}$,

$$\int_c^d f(x)\, dx = \int_{\varphi^{-1}(c)}^{\varphi^{-1}(d)} \big(f \circ \varphi(t)\big)\varphi'(t)\, dt.$$

Proof Set $F(t) = \int_{\varphi(a)}^{\varphi(t)} f(x)\, dx$. Then, $F(a) = 0$ and, by the chain rule and Theorem 3.2.1, $F' = (f \circ \varphi)\varphi'$. Hence, again by Theorem 3.2.1,

$$F(b) = \int_a^b \big(f \circ \varphi(t)\big)\varphi'(t)\, dt.$$

□

The equation in Corollary 3.2.3 is called the *change of variables formula*. In applications one often uses the mnemonic device of writing $x = \varphi(t)$ and $dx = \varphi'(t)\, dt$. For example, consider the integral $\int_0^1 \sqrt{1 - x^2}\, dx$, and make the change of variables $x = \sin t$. Then $dx = \cos t\, dt$, $0 = \arcsin 0$, and $\frac{\pi}{2} = \arcsin 1$. Hence $\int_0^1 \sqrt{1 - x^2}\, dx = \int_0^{\frac{\pi}{2}} \cos^2 t\, dt$, which, as we saw in connection with Wallis's formula, equals $\frac{\pi}{4}$. In that $\left\{\left(x, \sqrt{1 - x^2}\right) : x \in [0, 1]\right\}$ is the portion of the unit circle in the first quadrant, this is the answer that one should have expected.

Here is a slightly more theoretical application of Theorem 3.2.1.

Corollary 3.2.4 *Suppose that $\{\varphi_n : n \ge 1\}$ is a sequence of $\mathbb{C}$-valued continuous functions on $[a, b]$ that are continuously differentiable on (a, b). Further, assume that $\{\varphi_n : n \ge 1\}$ converges uniformly on $[a, b]$ to a function φ and that there is a function ψ on (a, b) such that $\varphi_n' \longrightarrow \psi$ uniformly on $[a + \delta, b - \delta]$ for each $0 < \delta < \frac{b-a}{2}$. Then φ is continuously differentiable on (a, b) and $\varphi' = \psi$.*

Proof Given $x \in (a, b)$, choose $0 < \delta < \frac{b-a}{2}$ so that $x \in (a+\delta, b-\delta)$. Then

$$\varphi(x) - \varphi(a+\delta) = \lim_{n\to\infty} \int_{a+\delta}^{x} \varphi_n'(y)\,dy = \int_{a+\delta}^{x} \psi(y)\,dy,$$

and so φ is differentiable at x and $\varphi'(x) = \psi(x)$. □

3.3 Rate of Convergence of Riemann Sums

In applications it is often very important to have an estimate on the rate at which Riemann sums are converging to the integral of a function. There are many results that deal with this question, and in this section we will show that there are circumstances in which the convergence is faster than one might have expected.

Given a continuous function $f : [0, 1] \longrightarrow \mathbb{C}$, it is obvious that

$$\left|\mathcal{R}(f; \mathcal{C}, \Xi) - \int_0^1 f(x)\,dx\right| \le \sup\bigl\{|f(y) - f(x)| : x,\ y \in [0,1] \text{ with } |y-x| \le \|\mathcal{C}\|\bigr\}$$

for any finite collection $\mathcal{C}$ of non-overlapping closed intervals whose union is $[0, 1]$ and any choice function Ξ. Hence, if f is continuously differentiable on $(0, 1)$, then

$$\left|\mathcal{R}(f; \mathcal{C}, \Xi) - \int_0^1 f(x)\,dx\right| \le \|f'\|_{(0,1)}\|\mathcal{C}\|.$$

Moreover, even if f has more derivatives, this is the best that one can say in general. On the other hand, as we will now show, one can sometimes do far better.

For $n \ge 1$, take $\mathcal{C}_n = \{I_{m,n} : 1 \le m \le n\}$, where $I_{m,n} = \left[\frac{m-1}{n}, \frac{m}{n}\right]$, $\Xi_n(I_{m,n}) = \frac{m}{n}$, and, for continuous $f : [0, 1] \longrightarrow \mathbb{C}$, set

$$\mathcal{R}_n(f) \equiv \mathcal{R}(f; \mathcal{C}_n, \Xi_n) = \frac{1}{n}\sum_{m=1}^{n} f\left(\tfrac{m}{n}\right).$$

Next, assume that f has a bounded, continuous derivative on $(0, 1)$, and apply integration by parts to each term, to see that

$$\int_0^1 f(x)\,dx - \mathcal{R}_n(f) = \sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(f(x) - f\left(\tfrac{m}{n}\right)\right) dx$$
$$= \sum_{m=1}^{\infty} \int_{I_{m,n}} \left(x - \tfrac{m-1}{n}\right)' \Bigl(f(x) - f\left(\tfrac{m}{n}\right)\Bigr)\,dx = -\sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \tfrac{m-1}{n}\right) f'(x)\,dx.$$

Now add the assumption that $f(1) = f(0)$. Then $\int_0^1 f'(x)\,dx = f(1) - f(0) = 0$, and so

$$\sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \tfrac{m-1}{n}\right) f'(x)\,dx = \sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \tfrac{m-1}{n} - c\right) f'(x)\,dx$$

for any constant c. In particular, by taking $c = \frac{1}{2n}$ to make each of the integrals $\int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \frac{m-1}{n} - c\right) dx = 0$, we can write

$$\int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \tfrac{m-1}{n} - \tfrac{1}{2n}\right) f'(x)\,dx = \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \tfrac{m-1}{n} - \tfrac{1}{2n}\right)\left(f'(x) - f'(\tfrac{m}{n})\right) dx,$$

and thereby conclude that

$$\int_0^1 f(x)\,dx - \mathcal{R}_n(f) = -\sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \tfrac{m-1}{n} - \tfrac{1}{2n}\right)\left(f'(x) - f'(\tfrac{m}{n})\right) dx.$$

This already represents progress. Indeed, because $\int_{\frac{m-1}{n}}^{\frac{m}{n}} \left|x - \frac{m-1}{n} - \frac{1}{2n}\right| dx = \frac{1}{4n^2}$ for each $1 \le m \le n$, we have shown that

$$\left|\int_0^1 f(x)\,dx - \mathcal{R}_n(f)\right| \le \frac{\sup\left\{|f'(y) - f'(x)| : |y - x| \le \frac{1}{n}\right\}}{4n}$$

if f is continuously differentiable and $f(1) = f(0)$. Hence, if in addition, f is twice differentiable, then $|f'(y) - f'(x)| \le \|f''\|_{(0,1)}|y - x|$ and the preceding leads to

$$(*) \qquad \left|\int_0^1 f(x)\,dx - \mathcal{R}_n(f)\right| \le \frac{\|f''\|_{(0,1)}}{4n^2}.$$

Before proceeding, it is important to realize how crucial a role the choice of both $\mathcal{C}_n$ and Ξ_n play. The role that $\mathcal{C}_n$ plays is reasonably clear, since it is what allowed us to choose the constant c independent of the intervals. The role of Ξ_n is more subtle. To see that it is essential, consider the function $f(x) = e^{i2\pi x}$, which is both smooth and satisfies $f(1) = f(0)$. Furthermore, $\|f'\|_{[0,1]} = 2\pi$ and $\int_0^1 f(x)\,dx = 0$. However, if $\tilde{\Xi}_n(I_{m,n}) = \frac{m(n-1)}{n^2}$, then

$$\mathcal{R}(f; \mathcal{C}_n, \tilde{\Xi}_n) = \frac{1}{n}\sum_{m=1}^{n} e^{i2\pi\frac{m(n-1)}{n^2}} = \frac{e^{i2\pi\frac{n-1}{n^2}}}{n}\,\frac{1 - e^{i2\pi\frac{n-1}{n}}}{1 - e^{i2\pi\frac{n-1}{n^2}}}.$$

Since $n\big(1-e^{i2\pi\frac{n-1}{n}}\big) = n\big(1-e^{-i2\pi\frac{1}{n}}\big) \longrightarrow i2\pi$ and $n\big(1-e^{i2\pi\frac{n-1}{n^2}}\big) \longrightarrow -i2\pi$, it follows that

$$\lim_{n\to\infty} n\left|\mathcal{R}(f;\mathcal{C}_n,\tilde{\Xi}_n) - \int_0^1 f(x)\,dx\right| = 1.$$

Thus, the analog of $(*)$ does not hold if one replaces Ξ_n by $\tilde{\Xi}_n$.

To go further, we introduce the notation

$$\Delta_n^{(k)}(f) \equiv \frac{1}{k!}\sum_{m=1}^{n}\int_{\frac{m-1}{n}}^{\frac{m}{n}}\big(x-\tfrac{m-1}{n}\big)^k\big(f(x)-f(\tfrac{m}{n})\big)\,dx \quad \text{for } k\ge 0.$$

Obviously, $\Delta_n^{(0)}(f) = \int_0^1 f(x)\,dx - \mathcal{R}_n(f)$. Furthermore, what we showed when f is continuously differentiable and $f(1)=f(0)$ is that $\Delta_n^{(0)}(f) = \frac{1}{2n}\Delta_n^{(0)}(f') - \Delta_n^{(1)}(f')$. By essentially the same argument, what we will now show is that, under the same assumptions on f,

$$\Delta_n^{(k)}(f) = \frac{1}{(k+2)!n^{k+1}}\Delta_n^{(0)}(f') - \Delta_n^{(k+1)}(f') \quad \text{for } k\ge 0. \tag{3.3.1}$$

The first step is to integrate each term by parts and thereby get

$$\Delta_n^{(k)}(f) = -\frac{1}{(k+1)!}\sum_{m=1}^{n}\int_{\frac{m-1}{n}}^{\frac{m}{n}}\big(x-\tfrac{m-1}{n}\big)^{k+1} f'(x)\,dx. \tag{3.3.2}$$

Because $\int_0^1 f'(x)\,dx = 0$, the right hand side does not change if one subtracts $\frac{1}{(k+2)n^{k+1}}$ from each $\big(x-\frac{m-1}{n}\big)^{k+1}$, and, once this subtraction is made, one can, with impunity, subtract $f'(\frac{m}{n})$ from $f'(x)$ in each term. In this way one arrives at

$$\Delta_n^{(k)}(f) = -\frac{1}{(k+1)!}\sum_{m=1}^{n}\int_{\frac{m-1}{n}}^{\frac{m}{n}}\Big(\big(x-\tfrac{m-1}{n}\big)^{k+1} - \tfrac{1}{(k+2)n^{k+1}}\Big)\big(f'(x)-f'(\tfrac{m}{n})\big)\,dx,$$

which, after rearrangement, is (3.3.1).

We will say that a function $\varphi:\mathbb{R}\longrightarrow\mathbb{C}$ is *periodic*[2] if $\varphi(x+1)=\varphi(x)$ for all $x\in\mathbb{R}$. Notice that if $\varphi:\mathbb{R}\longrightarrow\mathbb{C}$ is a bounded periodic function that is Riemann integrable on bounded intervals, then, by (3.1.6),

$$\int_a^{a+1}\varphi(\xi)\,d\xi = \int_0^1 \varphi(\xi)\,d\xi \quad \text{for all } a\in\mathbb{R}. \tag{3.3.3}$$

[2]In general, a function $f:\mathbb{R}\longrightarrow\mathbb{C}$ is said to be periodic if there is some $\alpha>0$ such that $f(x+\alpha)=f(x)$ for all $x\in\mathbb{R}$, in which case α is said to be a period of f. Here, without further comment, we will always be dealing with the case when $\alpha=1$ unless some other period is specified.

To check this, suppose that $n \le a < n + 1$. Then, by (3.1.6),

$$\begin{aligned}\int_a^{a+1} \varphi(\xi)\,d\xi &= \int_a^{n+1} \varphi(\xi)\,d\xi + \int_{n+1}^{a+1} \varphi(\xi)\,d\xi \\ &= \int_a^{n+1} \varphi(\xi)\,d\xi + \int_n^{a} \varphi(\xi)\,\xi = \int_n^{n+1} \varphi(\xi)\,d\xi = \int_0^{1} \varphi(\xi)\,d\xi.\end{aligned}$$

Now assume that $f : \mathbb{R} \longrightarrow \mathbb{C}$ is periodic and has $\ell \ge 1$ continuous derivatives. Starting from (3.3.1) and working by induction on ℓ, we see that

$$\Delta_n^{(0)}(f) = \frac{1}{n^\ell}\sum_{k=0}^{\ell} a_{k,\ell} n^{k+1} \Delta_n^{(k)}(f^{(\ell)}),$$

where

$$a_{0,0} = 1, \quad a_{0,\ell+1} = \sum_{k=0}^{\ell} \frac{a_{k,\ell}}{(k+2)!}, \quad \text{and } a_{k,\ell+1} = -a_{k-1,\ell} \text{ for } 1 \le k \le \ell.$$

The preceding can be simplified by observing that $a_{k,\ell} = (-1)^k a_{0,\ell-k}$ for $0 \le k \le \ell$, which means that

$$\begin{gathered}\Delta_n^{(0)}(f) = \frac{1}{n^{\ell+1}}\sum_{k=0}^{\ell} (-1)^k b_{\ell-k} n^{k+1} \Delta_n^{(k)}(f^{(\ell)}) \\ \text{where } b_0 = 1 \text{ and } b_{\ell+1} = \sum_{k=0}^{\ell} \frac{(-1)^k}{(k+2)!} b_{\ell-k}.\end{gathered} \tag{3.3.4}$$

If we now assume that f is $(\ell+1)$ times continuously differentiable, then, by (3.3.2),

$$\left|\Delta_n^{(k)}(f^{(\ell)})\right| \le \frac{1}{(k+1)!}\sum_{m=1}^{n} \int_{\frac{m-1}{n}}^{\frac{m}{n}} \left(x - \tfrac{m-1}{n}\right)^{k+1} |f^{(\ell+1)}(x)|\,dx \le \frac{\|f^{(\ell+1)}\|_{[0,1]}}{(k+2)!n^{k+1}},$$

and so (3.3.4) leads to the estimate (see Exercise 3.13 for a related estimate)

$$\left|\int_0^1 f(x)\,dx - \mathcal{R}_n(f)\right| \le \frac{K_\ell \|f^{(\ell+1)}\|_{[0,1]}}{n^{\ell+1}} \quad \text{where } K_\ell = \sum_{k=0}^{\ell} \frac{|b_{\ell-k}|}{(k+2)!} \tag{3.3.5}$$

for periodic functions f having $(\ell+1)$ continuous derivatives. In other words, if f is periodic and has $(\ell+1)$ continuous derivatives, then the Riemann sum $\mathcal{R}_n(f)$ differs from $\int_0^1 f(x)\,dx$ by at most the constant K_ℓ times $\|f^{(\ell+1)}\|_{[0,1]} n^{-\ell-1}$.

To get a feeling for how K_ℓ grows with ℓ, begin by taking $f(x) = e^{i2\pi x}$. Then $\Delta_1^{(0)}(f) = -1$, $\|f^{(\ell+1)}\|_{[0,1]} = (2\pi)^{\ell+1}$, and so (3.3.5) says that $K_\ell \ge (2\pi)^{-\ell-1}$. To get a crude upper bound, let α be the unique element of $(0, 1)$ for which $e^{\frac{1}{\alpha}} = 1 + \frac{2}{\alpha}$, and use induction on ℓ to see that $|b_\ell| \leq \alpha^\ell$. Hence, we now know that

$$(2\pi)^{-\ell-1} \leq K_\ell \leq e^{\frac{1}{\alpha}} \alpha^{\ell+2}.$$

Below we will get a much more precise result (cf. (3.4.9)), but in the meantime it should be clear that (3.3.5) is saying that the convergence of $\mathcal{R}_n(f)$ to $\int_0^1 f(x)\,dx$ is very fast when f is periodic, smooth, and the successive derivatives of f are growing at a moderate rate.

3.4 Fourier Series

Taylor's Theorem provides a systematic method for finding polynomial approximations to a function by scrutinizing in great detail the behavior of the function at a point. Although the method has many applications, it also has flaws. In fact, as we saw in the discussion following Lemma 1.8.3, there are circumstances in which it yields no useful information. As that example shows, the problem is that the behavior of a function away from a point cannot always be predicted from the behavior of it and its derivatives at the point. Speaking metaphorically, Taylor's method is analogous to attempting to lift an entire plank from one end.

Fourier introduced a very different approximation procedure, one in which the approximation is in terms of waves rather than polynomials. He took the trigonometric functions $\{\cos(2\pi mx) : m \geq 0\}$ and $\{\sin(2\pi mx) : m \geq 1\}$ as the model waves into which he wanted to resolve other functions. That is, he wanted to write a general continuous function $f : [0, 1] \longrightarrow \mathbb{C}$ as a usually infinite linear combination of the form

$$f(x) = \sum_{m=0}^{\infty} a_m \cos(2\pi mx) + \sum_{m=1}^{\infty} b_m \sin(2\pi mx),$$

where the coefficients a_m and b_m are complex numbers. Using (2.2.1), one sees that by taking $\hat{f}_0 = a_0$, $\hat{f}_m = \frac{a_m - ib_m}{2}$ for $m \geq 1$, and $\hat{f}_m = \frac{a_{-m} + ib_{-m}}{2}$ for $m \leq -1$, an equivalent and more convenient expression is

$$f(x) = \sum_{m=-\infty}^{\infty} \hat{f}_m \mathfrak{e}_m(x) \quad \text{where } \mathfrak{e}_m(x) \equiv e^{i2\pi mx}. \tag{3.4.1}$$

One of Fourier's key observations is that if one assumes that f can be represented in this way and that the series converges well enough, then the coefficients $\hat{f}_m$ are given by

$$\hat{f}_m \equiv \int_0^1 f(x)\mathfrak{e}_{-m}(x)\,dx. \tag{3.4.2}$$

To see this, observe that (cf. Exercise 3.7 for an alternative method)

$$\begin{aligned}
&\int_0^1 \mathfrak{e}_m(x)\mathfrak{e}_{-n}(x)\,dx \\
&\quad= \int_0^1 \big(\cos(2\pi mx)\cos(2\pi nx) + \sin(2\pi mx)\sin(2\pi nx)\big)\,dx \\
&\qquad\qquad + i\int_0^1 \big(-\cos(2\pi mx)\sin(2\pi nx) + \sin(2\pi mx)\cos(2\pi nx)\big)\,dx \\
&\quad= \int_0^1 \cos\big(2\pi(m-n)x\big)\,dx + i\int_0^1 \sin\big(2\pi(m-n)x\big)\,dx = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n, \end{cases}
\end{aligned}$$

where, in the passage to the last line, we used (1.5.1). Hence, if the exchange in the order of summation and integration is justified, (3.4.2) follows from (3.4.1).

We now turn to the problem of justifying Fourier's idea. That is, if $\hat{f}_m$ is given by (3.4.2), we want to examine to what extent it is true that f is represented by (3.4.1). Thus let a continuous $f : [0, 1] \longrightarrow \mathbb{C}$ be given, determine $\hat{f}_m$ accordingly by (3.4.2), and define

$$f_r(x) = \sum_{m=-\infty}^{\infty} r^{|m|}\hat{f}_m \mathfrak{e}_m(x) \quad \text{for } r \in [0, 1) \text{ and } x \in \mathbb{R}. \tag{3.4.3}$$

Because $|\hat{f}_m| \leq \|f\|_{[0,1]}$, the series defining f_r is absolutely convergent. In order to understand what happens when $r \nearrow 1$, observe that

$$\begin{aligned}
\sum_{m=-n}^{n} r^{|m|}\hat{f}_m \mathfrak{e}_m(x) &= \int_0^1 \left(\sum_{m=0}^{n} \big(r\mathfrak{e}_1(x-y)\big)^m\right) f(y)\,dy \\
&\qquad + \int_0^1 \left(\sum_{m=1}^{n} \big(r\mathfrak{e}_1(y-x)\big)^m\right) f(y)\,dy \\
&= \int_0^1 \left(\frac{1 - r^{n+1}\mathfrak{e}_{n+1}(x-y)}{1 - r\mathfrak{e}_1(x-y)} + \frac{r\mathfrak{e}_1(y-x) - r^{n+1}\mathfrak{e}_{n+1}(y-x)}{1 - r\mathfrak{e}_1(y-x)}\right) f(y)dy \\
&= \int_0^1 \frac{1 - r^2 - r^{n+1}\cos\big(2\pi(n+1)(x-y)\big) + r^{n+2}\cos(2\pi n(x-y))}{|1 - r\mathfrak{e}_1(x-y)|^2} f(y)\,dy.
\end{aligned}$$

Hence, by Theorem 3.1.4,

$$f_r(x) = \sum_{m=-\infty}^{\infty} r^{|m|}\hat{f}_m \mathfrak{e}_m(x) = \int_0^1 p_r(x-y)f(y)\,dy$$

$$\text{where } p_r(\xi) \equiv \frac{1 - r^2}{|1 - r\mathfrak{e}_1(\xi)|^2} \text{ for } \xi \in \mathbb{R}. \tag{3.4.4}$$

Obviously the function p_r is positive, periodic, and even: $p_r(-\xi) = p_r(\xi)$. Furthermore, by taking $f = \mathbf{1}$ and noting that then $\hat{f}_0 = 1$ and $\hat{f}_m = 0$ for $m \neq 0$, we see from (3.4.4) and evenness that $\int_0^1 p_r(x-y)\,dy = 1$. Finally, note that if $\delta \in \big(0, \frac{1}{2}\big)$ and $\xi \notin \bigcup_{n=-\infty}^{\infty}[n-\delta, n+\delta]$, then

$$(*) \qquad |1 - r\mathfrak{e}_1(\xi)|^2 \geq 2\big(1-\cos(2\pi\xi)\big) \geq \omega(\delta) \equiv 2\big(1-\cos(2\pi\delta)\big) > 0$$

and therefore $p_r(\xi) \leq \frac{1-r^2}{\omega(\delta)}$.

Given a function $f : [0,1] \longrightarrow \mathbb{C}$, its *periodic extension* to $\mathbb{R}$ is the function $\tilde{f}$ given by $\tilde{f}(x) = f(x-n)$ for $n \leq x < n+1$. Clearly, $\tilde{f}$ will be continuous if and only if f is continuous and $f(0) = f(1)$. In addition, if $\ell \geq 1$, then $\tilde{f}$ will be ℓ times continuously differentiable if and only if f is ℓ times continuously differentiable on $(0,1)$ and the limits $\lim_{x\searrow 0} f^{(k)}(x)$ and $\lim_{x\nearrow 1} f^{(k)}(x)$ exist and are equal for each $0 \leq k \leq \ell$.

Theorem 3.4.1 *Let $f : [0,1] \longrightarrow \mathbb{C}$ be a continuous function, and define f_r for $r \in [0,1)$ as in (3.4.3). Then, for each $r \in [0,1)$, f_r is a periodic function with continuous derivatives of all orders. In fact, for each $k \geq 0$, the series $\sum_{m=-\infty}^{\infty}(i2\pi|m|)^k r^{|m|}\hat{f}_m\mathfrak{e}_m(x)$ is absolutely and uniformly convergent to $f_r^{(k)}(x)$. Furthermore, for each $\delta \in \big(0,\frac{1}{2}\big)$, $f_r \longrightarrow f$ as $r \nearrow 1$ uniformly on $[\delta, 1-\delta]$. Finally, if $f(0) = f(1)$, then $f_r \longrightarrow f$ uniformly on $[0,1]$.*

Proof Since each term in the sum defining f_r is periodic, it is obvious that f_r is also. In addition, since the sum converges uniformly on $[0,1]$, f_r is continuous. To prove that f_r has continuous derivatives of all orders, we begin by applying Theorem 2.3.2 to see that $\sum_{m=-\infty}^{\infty}|m|^k r^{|m|} \leq 2\sum_{m=0}^{\infty} m^k r^m < \infty$.

In view of the preceding, we know that $\sum_{m=-\infty}^{\infty}(i2\pi|m|)r^{|m|}\hat{f}_m\mathfrak{e}_m(x)$ converges uniformly for $x \in [0,1]$ to a function ψ. At the same time, if $\varphi_n(x) \equiv \sum_{m=-n}^{n} r^{|m|}\hat{f}e_m(x)$, then φ_n converges uniformly to f_r and $\varphi_n' = \sum_{m=-n}^{n}(i2\pi m)r^{|m|}\hat{f}_m\mathfrak{e}_m(x)$ converges uniformly to ψ. Hence, by Corollary 3.2.4, f_r is differentiable and $f_r' = \psi$. More generally, assume that f_r has k continuous derivatives and that $f_r^{(k)} = \sum_{m=-\infty}^{\infty} r^{|m|}(i2\pi|m|)^k\hat{f}_m\mathfrak{e}_m$. Then a repetition of the preceding argument shows that $f_r^{(k)}$ is differentiable and that its derivative is $\sum_{m=-\infty}^{\infty} r^{|m|}(i2\pi|m|)^{k+1}\hat{f}_m\mathfrak{e}_m$.

Now take $\tilde{f}$ to be the periodic extension of f to $\mathbb{R}$. Since $\tilde{f}$ is bounded and Riemann integrable on bounded intervals, (3.3.3) together with $\int_0^1 p_r(x-y)\,dy = 1$ show that

$$\begin{aligned} f_r(x) - f(x) &= \int_0^1 p_r(x-y)\big(f(y)-f(x)\big)\,dy \\ &= \int_{-x}^{1-x} p_r(\xi)\big(f(\xi+x)-f(x)\big)\,d\xi = \int_{-\frac{1}{2}}^{\frac{1}{2}} p_r(\xi)\big(\tilde{f}(\xi+x)-\tilde{f}(x)\big)\,d\xi \end{aligned}$$

$$= \int_{-\frac{1}{2}}^{-\delta} p_r(\xi)\big(\tilde{f}(\xi+x) - \tilde{f}(x)\big)\,d\xi + \int_{-\delta}^{\delta} p_r(\xi)\big(\tilde{f}(\xi+x) - \tilde{f}(x)\big)\,d\xi$$
$$+ \int_{\delta}^{\frac{1}{2}} p_r(\xi)\big(\tilde{f}(\xi+x) - \tilde{f}(x)\big)\,d\xi$$

for $x \in [0, 1]$ and $0 < \delta < \frac{1}{2}$. Using $(*)$, one sees that the first and third terms in the final expression are dominated by $2\|f\|_{[0,1]}\frac{1-r^2}{\omega(\delta)}$ and therefore tend to 0 as $r \nearrow 1$. As for the second term, it is dominated by $\sup\{|\tilde{f}(y) - \tilde{f}(x)| : |y-x| \le \delta\}$. Hence, if $\tilde{f}$ is continuous at x, as it will be if $x \in (0, 1)$, then

$$\lim_{r\nearrow 1}\big|f_r(x) - f(x)\big| = \lim_{r\nearrow 1}\left|\int_0^1 p_r(x-y)\big(f(y) - f(x)\big)\,dy\right| = 0.$$

Moreover, the convergence is uniform on the interval $[\delta, 1-\delta]$, and if $f(0) = f(1)$ and therefore $\tilde{f}$ is continuous everywhere, then the convergence is uniform on $[0, 1]$. □

Even though the preceding result is a weakened version of it, we now know that Fourier's idea is basically sound. One important fact to which this weak version leads is the identity

$$\int_0^1 f(x)\overline{g(x)}\,dx = \sum_{m=-\infty}^{\infty} \hat{f}_m\overline{\hat{g}_m}. \tag{3.4.5}$$

To prove this, first observe that, because $\int_0^1 \mathfrak{e}_{m_1}(x)\mathfrak{e}_{-m_2}(x)\,dx$ is 0 if $m_1 \ne m_2$ and 1 if $m_1 = m_2$, one can use Theorem 3.1.4 to justify

$$\int_0^1 f_r(x)\overline{g_r(x)}\,dx = \sum_{m_1,m_2=-\infty}^{\infty} r^{|m_1|+|m_2|}\hat{f}_{m_1}\overline{\hat{g}_{m_2}}\int_0^1\int_0^1 \mathfrak{e}_{m_1}(x)\mathfrak{e}_{-m_2}(x)\,dx$$
$$= \sum_{m=-\infty}^{\infty} r^{2|m|}\hat{f}_m\overline{\hat{g}_m}.$$

At the same time, for each $\delta \in \big(0, \frac{1}{2}\big)$,

$$\varlimsup_{r\nearrow 1}\int_0^1\big|f_r(x)\overline{g_r(x)} - f(x)\overline{g(x)}\big|\,dx$$
$$\le \varlimsup_{r\nearrow 1}\int_0^{\delta}\big|f_r(x)\overline{g_r(x)} - f(x)\overline{g(x)}\big|\,dx + \varlimsup_{r\nearrow 1}\int_{1-\delta}^{1}\big|f_r(x)\overline{g_r(x)} - f(x)\overline{g(x)}\big|\,dx$$
$$\le 4\|f\|_{[0,1]}\|g\|_{[0,1]}\delta,$$

and so $\lim_{r\nearrow 1}\int_0^1 f_r(x)\overline{g_r(x)}\,dx = \int_0^1 f(x)\overline{g(x)}\,dx$. Taking $g = f$, we conclude that

$$\int_0^1 |f(x)|^2\,dx = \lim_{r\nearrow 1}\sum_{m=-\infty}^{\infty} r^{2|m|}|\hat{f}_m|^2,$$

and therefore that $\sum_{m=-\infty}^{\infty}|\hat{f}_m|^2 < \infty$. Hence, by Schwarz's inequality (cf. Exercise 2.3),

$$\sum_{m=-\infty}^{\infty}|\hat{f}_m||\hat{g}_m| \le \sqrt{\sum_{m=-\infty}^{\infty}|\hat{f}_m|^2}\sqrt{\sum_{m=-\infty}^{\infty}|\hat{g}_m|^2} < \infty,$$

and so the series $\sum_{m=-\infty}^{\infty}\hat{f}_m\overline{\hat{g}_m}$ is absolutely convergent. Thus, when we apply (1.10.1), we find that

$$\sum_{m=-\infty}^{\infty}\hat{f}_m\overline{\hat{g}_m} = \lim_{r\nearrow 1}\sum_{m=-\infty}^{\infty} r^{2|m|}\hat{f}_m\overline{\hat{g}_m} = \int_0^1 f(x)\overline{g(x)}\,dx.$$

The identity in (3.4.5) is known as *Parseval's equality*, and it has many interesting applications of which the following is an example. Take $f(x) = x$. Obviously, $\hat{f}_0 = \frac{1}{2}$, and, using integration by parts, one sees that $\hat{f}_m = \frac{1}{i2\pi m}$ for $m \neq 0$. Hence, by (3.4.5),

$$\frac{1}{3} = \int_0^1 |f(x)|^2\,dx = \frac{1}{4} + \frac{1}{4\pi^2}\sum_{m\neq 0}\frac{1}{m^2} = \frac{1}{4} + \frac{1}{2\pi^2}\sum_{m=1}^{\infty}\frac{1}{m^2},$$

from which we see that

$$\sum_{m=1}^{\infty}\frac{1}{m^2} = \frac{\pi^2}{6}. \tag{3.4.6}$$

The function $\zeta(z)$ given by $\sum_{m=1}^{\infty}\frac{1}{m^z}$ when the real part of z is greater than 1 is the famous *Riemann zeta function* which plays an important role in number theory (cf. Sect. 6.3). We now know the value of ζ at $z = 2$, and, as we will show below, we can also find its value at all positive, even integers. However, in order to do that computation, we will need to discuss when the *Fourier series* $\sum_{m=-\infty}^{\infty}\hat{f}_m\mathfrak{e}_m(x)$ converges to f. This turns out to be a very delicate question, and we will not attempt to describe any of the refined answers that have been found. Instead, we will deal only with the most elementary case.

Theorem 3.4.2 *Let $f : [0,1] \longrightarrow \mathbb{C}$ be a continuous function. If $\sum_{m=-\infty}^{\infty}|\hat{f}_m| < \infty$, then $f(0) = f(1)$ and $\sum_{m=-\infty}^{\infty}\hat{f}_m\mathfrak{e}_m$ converges uniformly to f. In particular, if $f(1) = f(0)$ and f has a bounded, continuous derivative on $(0,1)$, then $\sum_{m=-\infty}^{\infty}|\hat{f}_m| < \infty$ and therefore $\sum_{m=-\infty}^{\infty}\hat{f}_m\mathfrak{e}_m$ converges uniformly to f.*

Proof We already know that $f_r(x) \longrightarrow f(x)$ for all $x \in (0,1)$, and therefore, by (1.10.1), if $x \in (0,1)$ and $\sum_{m=-\infty}^{\infty} \hat{f}_m \mathfrak{e}_m(x)$ converges, it must converge to $f(x)$. Thus, since $|\hat{f}_m \mathfrak{e}_m(x)| \le |\hat{f}_m|$ if $\sum_{m=-\infty}^{\infty} |\hat{f}_m| < \infty$, $\sum_{m=-\infty}^{\infty} \hat{f}_m \mathfrak{e}_m(x)$ converges absolutely and uniformly on [0, 1] to a continuous, periodic function, and, because that function coincides with f on (0, 1), it must be equal to f on [0, 1].

Now assume that $f(1) = f(0)$ is and that f has a bounded, continuous derivative f' on (0, 1). Using integration by parts and Exercise 3.7, we see that, for $m \neq 0$,

$$
\begin{aligned}
\hat{f}_m &= \lim_{\delta \searrow 0} \int_\delta^{1-\delta} f(x)\mathfrak{e}_{-m}(x)\,dx \\
&= \lim_{\delta \searrow 0} \frac{1}{i2\pi m}\left(f(\delta)\mathfrak{e}_{-m}(\delta) - f(1-\delta)\mathfrak{e}_{-m}(1-\delta) + \int_\delta^{1-\delta} f'(x)\mathfrak{e}_{-m}(x)\,dx \right) \\
&= \frac{\widehat{f'_m}}{i2\pi m}.
\end{aligned}
$$

Hence, by Schwarz's inequality, Parseval's equality, and (3.4.6),

$$
\sum_{m \neq 1} |\hat{f}_m| \le \frac{1}{2\pi}\left(\sum_{m\neq 0}\frac{1}{m^2}\right)^{\frac{1}{2}} \left(\sum_{m \neq 0} |\widehat{f'_m}|^2\right)^{\frac{1}{2}} \le \sqrt{\frac{1}{24}}\left(\int_0^1 |f'(x)|^2\,dx\right)^{\frac{1}{2}}. \qquad \square
$$

Notice that the integration by parts step in the preceding has the following easy extension. Namely, suppose that $f : [0,1] \longrightarrow \mathbb{C}$ is continuous and that f has $\ell \ge 1$ bounded, continuous derivatives on (0, 1). Further, assume that $f^{(k)}(0) \equiv \lim_{x \searrow 0} f^{(k)}(x)$ and $f^{(k)}(1) \equiv \lim_{x \nearrow 1} f(x)$ exist and are equal for $0 \le k < \ell$. Then, by iterating the argument given above, one sees that

$$
\hat{f}_m = (i2\pi m)^{-\ell}\widehat{(f^{(\ell)})}_m \quad \text{for } m \in \mathbb{Z} \setminus \{0\}.
$$

Returning to the computation of $\zeta(2\ell)$, recall the numbers b_k introduced in (3.3.4), and set

$$
P_\ell(x) = \sum_{k=0}^{\ell} \frac{(-1)^k b_{\ell-k}}{k!} x^k \quad \text{for } \ell \ge 0 \text{ and } x \in \mathbb{R}.
$$

Then $P_0 \equiv 1$ and $P_\ell = -P'_{\ell+1}$. In addition, if $\ell \ge 2$, then

$$
P_\ell(1) = b_\ell - b_{\ell-1} + \sum_{k=2}^{\ell} \frac{(-1)^k b_{\ell-k}}{k!}
$$

and, by (3.3.4),

$$\sum_{k=2}^{\ell} \frac{(-1)^k b_{\ell-k}}{k!} = \sum_{k=0}^{\ell-2} \frac{(-1)^k b_{\ell-2-k}}{(k+2)!} = b_{\ell-1}.$$

Hence $P_\ell(1) = b_\ell = P_\ell(0)$ for all $\ell \geq 2$. In particular, these lead to

$$(\widehat{P}_\ell)_0 = -\int_0^1 P'_{\ell+1}(x)\,dx = P_{\ell+1}(0) - P_{\ell+1}(1) = 0 \quad \text{for} \quad \ell \geq 1$$

and

$$(\widehat{P}_\ell)_m = (-1)^{\ell-1}(i2\pi m)^{1-\ell}(\widehat{P}_1)_m \quad \text{for} \quad \ell \geq 2 \quad \text{and} \quad m \neq 0.$$

Since $P_1(x) = b_1 - b_0 x = \frac{1}{2} - x$, we can use integration by parts to see that

$$(\widehat{P}_1)_m = -\int_0^1 x e^{-i2\pi m x}\,dx = (i2\pi m)^{-1} \quad \text{for} \quad m \neq 0,$$

and therefore

$$P_\ell(x) = -\left(\frac{i}{2\pi}\right)^\ell \sum_{m \neq 0} \frac{\mathfrak{e}_m(x)}{m^\ell} \quad \text{for} \quad \ell \geq 2 \quad \text{and} \quad x \in [0,1]. \tag{3.4.7}$$

Taking $x = 0$ in (3.4.7), we have that

$$b_{2\ell+1} = 0 \quad \text{and} \quad b_{2\ell} = \frac{(-1)^{\ell+1} 2\zeta(2\ell)}{(2\pi)^{2\ell}} \tag{3.4.8}$$

for $\ell \geq 1$. Knowing that $b_\ell = 0$ for odd $\ell \geq 3$, the recursion relation in (3.3.4) for b_ℓ simplifies to

$$b_\ell = \begin{cases} 2^{-\ell} & \text{if } \ell \in \{0,1\} \\ \frac{1}{2\ell!} - \sum_{k=0}^{\frac{\ell-2}{2}} \frac{b_{2k}}{(\ell-2k+1)!} & \text{if } \ell \ \geq 2 \text{ is even} \\ 0 & \text{if } \ell \geq 2 \text{ is odd.} \end{cases} \tag{3.4.9}$$

Now we can go the other direction and use (3.4.8) and (3.4.9) to compute ζ at even integers:

$$\zeta(2\ell) = (-1)^{\ell+1} 2^{2\ell-1} \pi^{2\ell} b_{2\ell} \quad \text{for } \ell \geq 1. \tag{3.4.10}$$

Finally, starting from (3.4.10), one sees that the K_ℓ's in (3.3.5) satisfy $\lim_{\ell\to\infty}(K_\ell)^{\frac{1}{\ell}} = (2\pi)^{-1}$.

In the literature, the numbers $\ell! b_\ell$ are called the *Bernoulli numbers*, and they have an interesting history. Using (3.4.9) together with (3.4.10), one recovers (3.4.6) and sees that $\zeta(4) = \frac{\pi^4}{90}$ and $\zeta(6) = \frac{\pi^6}{945}$. Using these relations to compute ζ at larger even integers is elementary but tedious. Perhaps more interesting than such computations is the observation that, when $\ell \geq 2$, $P_\ell(x)$ is an ℓth order polynomial whose periodic extension from $[0, 1]$ is $(\ell - 2)$ times differentiable. That such polynomials exist is not obvious.

3.5 Riemann–Stieltjes Integration

The topic of this concluding section is an easy but important generalization, due to Stieltjes, of Riemann integration. Namely, given bounded, $\mathbb{R}$-valued functions φ and ψ on $[a, b]$, a finite cover $\mathcal{C}$ of $[a, b]$ by non-overlapping closed intervals I, and an associated choice function Ξ, set

$$\mathfrak{R}(\varphi|\psi; \mathcal{C}, \Xi) = \sum_{I \in \mathcal{C}} \varphi\big(\Xi(I)\big) \Delta_I \psi,$$

where $\Delta_I \psi$ denotes the difference between the value of ψ at the right hand end point of I and its value at the left hand end point. We will say that φ is *Riemann–Stieltjes integrable* on $[a, b]$ with respect to ψ if there exists a number $\int_a^b \varphi(x)\, d\psi(x)$, known as the *Riemann–Stieltjes* of φ with respect to ψ, such that for each $\epsilon > 0$ there is a $\delta > 0$ for which

$$\left| \mathfrak{R}(\varphi|\psi; \mathcal{C}, \Xi) - \int_a^b \varphi(x)\, d\psi(x) \right| < \epsilon$$

whenever $\|\mathcal{C}\| < \delta$ and Ξ is any associated choice function for $\mathcal{C}$. Obviously, when $\psi(x) = x$, this is just Riemann integration. In addition, it is clear that if φ_1 and φ_2 are Riemann–Stieltjes integrable with respect to ψ, then, for all $\alpha_1, \alpha_2 \in \mathbb{R}$, $\alpha_1\varphi_1 + \alpha_2\varphi_2$ is also and

$$\int_a^b \big(\alpha_1\varphi_1(x) + \alpha_2\varphi_2(x)\big)\, d\psi(x) = \alpha_1 \int_a^b \varphi_1(x)\, d\psi(x) + \alpha_2 \int_a^b \varphi_2(x)\, d\psi(x).$$

Also, if $\check{\psi}(x) = \psi(a + b - x)$ and φ is Riemann–Stieltjes integrable with respect to ψ, then $x \rightsquigarrow \varphi(a + b - x)$ is Riemann–Stieltjes integrable with respect to $\check{\psi}$ and

$$\int_a^b \varphi(a + b - x)\, d\check{\psi}(x) = -\int_a^b \varphi(x)\, d\psi(x). \tag{3.5.1}$$

In general it is hard to determine which functions φ are Riemann–Stieltjes integrable with respect to a given function ψ. Nonetheless, the following simple lemma shows that there is an inherent symmetry between the roles of φ and ψ.

Lemma 3.5.1 *If φ is Riemann–Stieltjes integrable with respect to ψ, then ψ is Riemann–Stieltjes integrable with respect to φ and*

$$\int_a^b \varphi(x)\,d\psi(x) = \varphi(b)\psi(b) - \varphi(a)\psi(b) - \int_a^b \psi(x)\,d\varphi(x). \tag{3.5.2}$$

Proof Let $\mathcal{C} = \{[\alpha_{m-1}, \alpha_m] : 1 \le m \le n\}$, where $a = \alpha_0 \le \cdots \le \alpha_n = b$, and let Ξ be an associated choice function. Set $\beta_0 = a$, $\beta_m = \Xi([\alpha_{m-1}, \alpha_m])$ for $1 \le m \le n$, and $\beta_{n+1} = b$, and define $\mathcal{C}' = \{[\beta_{m-1}, \beta_m] : 1 \le m \le n+1\}$ and $\Xi'([\beta_{m-1}, \beta_m]) = \alpha_{m-1}$ for $1 \le m \le n+1$. Then

$$\begin{aligned}
\mathcal{R}(\psi|\varphi; \mathcal{C}, \Xi) &= \sum_{m=1}^{n} \psi(\beta_m)\big(\varphi(\alpha_m) - \varphi(\alpha_{m-1})\big) \\
&= \sum_{m=1}^{n} \psi(\beta_m)\varphi(\alpha_m) - \sum_{m=0}^{n-1} \psi(\beta_{m+1})\varphi(\alpha_m) \\
&= \psi(\beta_n)\varphi(\alpha_n) - \sum_{m=1}^{n-1} \varphi(\alpha_m)\big(\psi(\beta_{m+1}) - \psi(\beta_m)\big) - \psi(\beta_1)\varphi(\alpha_0) \\
&= \psi(b)\varphi(b) - \psi(a)\varphi(a) - \sum_{m=0}^{n} \varphi(\alpha_m)\big(\psi(\beta_{m+1}) - \psi(\beta_m)\big) \\
&= \psi(b)\varphi(b) - \psi(a)\varphi(a) - \mathcal{R}(\varphi|\psi; \mathcal{C}', \Xi').
\end{aligned}$$

Noting that $\|\mathcal{C}'\| \le 2\|\mathcal{C}\|$, one now sees that if φ is Riemann–Stieltjes integrable with respect to ψ, then ψ is Riemann–Stieltjes integrable with respect to φ and (3.5.2) holds. □

As we will see, Lemma 3.5.1 is an interesting generalization of the integration by parts, but it does little to advance us toward an understanding of the basic problem. In addressing that problem, the following analog of Theorem 3.1.2 will play a central role.

Lemma 3.5.2 *If ψ is non-decreasing on $[a, b]$, then φ is Riemann–Stieltjes integrable with respect to ψ if and only if for each $\epsilon > 0$ there exists a $\delta > 0$ such that*

$$\|\mathcal{C}\| < \delta \implies \sum_{\substack{I \in \mathcal{C} \\ \sup_I \varphi - \inf_I \varphi > \epsilon}} \Delta_I \psi < \epsilon.$$

In particular, every continuous function φ is Riemann–Stieltjes integrable with respect to ψ. In addition, if φ is Riemann–Stieltjes integrable with respect to ψ and $c \in (a, b)$, then it is Riemann–Stieltjes integrable with respect of ψ on both $[a, c]$ and $[c, b]$ and

$$\int_a^b \varphi(x)\, d\psi(x) = \int_a^c \varphi(x)\, d\psi(x) + \int_c^b \varphi(x)\, d\psi(x).$$

Finally, if $\varphi : [a, b] \longrightarrow [c, d]$ is Riemann–Stieltjes integrable with respect to ψ and $f : [c, d] \longrightarrow \mathbb{R}$ is continuous, then $f \circ \varphi$ is again Riemann–Stieltjes integrable with respect to ψ.

Proof Once the first part is proved, the other assertions follow in exactly the same way as the analogous assertions followed from Theorem 3.1.2.

In proving the first part, we will assume, without loss in generality, that $\Delta \equiv \Delta_{[a,b]}\psi > 0$. Now suppose that φ is Riemann–Stieltjes integrable with respect to ψ. Given $\epsilon > 0$, choose $\delta > 0$ so that

$$\|\mathcal{C}\| \le \delta \implies \left|\mathfrak{R}(\varphi|\psi; \mathcal{C}, \Xi) - \int_a^b \varphi(x)\, d\psi(x)\right| \le \frac{\epsilon^2}{4}$$

for all associated choice functions Ξ. Next, given $\mathcal{C}$, choose, for each $I \in \mathcal{C}$, $\Xi_1(I),\ \Xi_2(I) \in I$ so that $\varphi\big(\Xi_1(I)\big) \ge \sup_I \varphi - \frac{\epsilon^2}{4\Delta}$ and $\varphi\big(\Xi_2(I)\big) \le \inf_I \varphi + \frac{\epsilon^2}{4\Delta}$. Then $\|\mathcal{C}\| < \delta$ implies that

$$\frac{\epsilon^2}{2} \ge \mathfrak{R}(\varphi|\psi; \mathcal{C}, \Xi_1) - \mathfrak{R}(\varphi|\psi; \mathcal{C}, \Xi_2) \ge \sum_{I \in \mathcal{C}} \left(\sup_I \varphi - \inf_I \varphi\right) \Delta_I \psi - \frac{\epsilon^2}{2},$$

and so

$$\epsilon^2 \ge \epsilon \sum_{\substack{I \in \mathcal{C} \\ \sup_I \varphi - \inf_I \varphi > \epsilon}} \Delta_I \psi.$$

To prove the converse, we introduce the upper and lower Riemann–Stieltjes sums

$$\mathcal{U}(\varphi|\psi; \mathcal{C}) = \sum_{I \in \mathcal{C}} \sup_I \varphi \Delta_I \psi \text{ and } \mathcal{L}(\varphi|\psi; \mathcal{C}) = \sum_{I \in \mathcal{C}} \inf_I \varphi \Delta_I \psi.$$

Just as in the Riemann case, one sees that

$$\mathcal{L}(\varphi|\psi; \mathcal{C}) \le \mathfrak{R}(\varphi|\psi; \mathcal{C}, \Xi) \le \mathcal{U}(\varphi|\psi; \mathcal{C})$$

for all $\mathcal{C}$ and associated rate functions Ξ, and $\mathcal{L}(\varphi|\psi; \mathcal{C}) \le \mathcal{U}(\varphi|\psi; \mathcal{C}')$ for all $\mathcal{C}$ and $\mathcal{C}'$. Further,

$$\mathcal{U}(\varphi|\psi;\mathcal{C}) - \mathcal{L}(\varphi|\psi;\mathcal{C}) = \sum_{I\in\mathcal{C}} \left(\sup_I \varphi - \inf_I \varphi\right) \Delta_I \psi$$
$$\le \epsilon\Delta + 2\|\varphi\|_{[a,b]} \sum_{\substack{I\in\mathcal{C} \\ \sup_I \varphi - \inf_I \varphi > \epsilon}} \Delta_I \psi,$$

and so, under the stated condition, for each $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\|\mathcal{C}\| < \delta \implies \mathcal{U}(\varphi|\psi;\mathcal{C}) - \mathcal{L}(\varphi|\psi;\mathcal{C}) < \epsilon.$$

As a consequence, we know that if $\|\mathcal{C}\| < \delta$ then, for any $\mathcal{C}'$

$$\mathcal{U}(\varphi|\psi;\mathcal{C}) \le \mathcal{L}(\varphi|\psi;\mathcal{C}) + \epsilon \le \mathcal{U}(\varphi|\psi;\mathcal{C}') + \epsilon,$$

and similarly, $\mathcal{L}(\varphi|\psi;\mathcal{C}) \ge \mathcal{L}(\varphi|\psi;\mathcal{C}') - \epsilon$. From these it follows that $M \equiv \inf_{\mathcal{C}} \mathcal{U}(\varphi|\psi;\mathcal{C}) = \sup_{\mathcal{C}} \mathcal{L}(\varphi|\psi;\mathcal{C})$ and

$$\lim_{\|\mathcal{C}\|\to 0} \mathcal{U}(\varphi|\psi;\mathcal{C}) = M = \lim_{\|\mathcal{C}\|\to 0} \mathcal{L}(\varphi|\psi;\mathcal{C}),$$

and at this point the rest of the argument is the same as the one in the proof of Theorem 3.1.2. □

The reader should take note of the distinction between the first assertion here and the analogous one in Theorem 3.1.2. Namely, in Theorem 3.1.2, the condition for Riemann integrability was that there exist some $\mathcal{C}$ for which

$$\sum_{\substack{I\in\mathcal{C} \\ \sup_I f - \inf_I f > \epsilon}} |I| < \epsilon,$$

whereas here we insist that

$$\|\mathcal{C}\| < \delta \implies \sum_{\substack{I\in\mathcal{C} \\ \sup_I \varphi - \inf_I \varphi > \epsilon}} \Delta_I \psi < \epsilon.$$

The reason for this is that the analog of the final assertion in Lemma 3.1.1 is true for ψ only if ψ is continuous. When ψ is continuous, then the condition that $\|\mathcal{C}\| < \delta$ can be removed from the first assertion in Lemma 3.5.2.

In order to deal with ψ's that are not monotone, we introduce the quantity

$$\mathrm{var}_{[a,b]}(\psi) \equiv \sup_{\mathcal{C}} \sum_{I\in\mathcal{C}} |\Delta_I \psi| < \infty,$$

where $\mathcal{C}$ denotes a generic finite cover of $[a, b]$ by non-overlapping closed intervals, and say that ψ has *finite variation* on $[a, b]$ if $\mathrm{var}_{[a,b]}(\psi) < \infty$. It is easy to check that

$\mathrm{var}_{[a,b]}(\psi_1 + \psi_2) \le \mathrm{var}_{[a,b]}(\psi_1) + \mathrm{var}_{[a,b]}(\psi_2)$ and that $\mathrm{var}_{[a,b]}(\psi) = |\psi(b) - \psi(a)|$ if ψ is monotone (i.e., it is either non-decreasing or non-increasing). Thus, if ψ can be written as the difference between two non-increasing functions ψ^+ and ψ^-, then it has bounded variation and

$$\mathrm{var}_{[a,b]}(\psi) \le \big(\psi^+(b) - \psi^+(a)\big) + \big(\psi^-(b) - \psi^-(a)\big).$$

We will now show that every function of bounded variation admits such a representation. To this end, define

$$\mathrm{var}^{\pm}_{[a,b]}(\psi) = \sup_{\mathcal{C}} \sum_{I \in \mathcal{C}} (\Delta_I \psi)^{\pm}.$$

Lemma 3.5.3 *If ψ has bounded variation on $[a, b]$, then*

$$\Delta_{[a,b]}\psi = \mathrm{var}^+_{[a,b]}(\psi) - \mathrm{var}^-_{[a,b]}(\psi) \textit{ and } \mathrm{var}_{[a,b]}(\psi) = \mathrm{var}^+_{[a,b]}(\psi) + \mathrm{var}^-_{[a,b]}(\psi).$$

Proof Obviously,

$$\sum_{I \in \mathcal{C}} |\Delta_I \psi| = \sum_{I \in \mathcal{C}} (\Delta_I \psi)^+ + \sum_{I \in \mathcal{C}} (\Delta_I \psi)^-$$

and

$$\Delta_{[a,b]}\psi = \sum_{I \in \mathcal{C}} (\Delta_I \psi)^+ - \sum_{I \in \mathcal{C}} (\Delta_I \psi)^-.$$

From the first of these, it is clear that $\mathrm{var}_{[a,b]}(\psi) \le \mathrm{var}^+_{[a,b]}(\psi) + \mathrm{var}^-_{[a,b]}(\psi)$. From the second we see that $\mathrm{var}^{\pm}_{[a,b]}(\psi) \le \mathrm{var}^{\mp}_{[a,b]}(\psi) \pm \Delta_{[a,b]}\psi$, and therefore that $\Delta_{[a,b]}\psi = \mathrm{var}^+_{[a,b]}(\psi) - \mathrm{var}^-_{[a,b]}(\psi)$. Hence, if $\lim_{n\to\infty} \sum_{I \in \mathcal{C}_n} (\Delta_I \psi)^+ = \mathrm{var}^+_{[a,b]}(\psi)$, then $\lim_{n\to\infty} \sum_{I \in \mathcal{C}_n} (\Delta_I \psi)^- = \mathrm{var}^-_{[a,b]}(\psi)$, and so

$$\mathrm{var}_{[a,b]}(\psi) \ge \lim_{n\to\infty} \left(\sum_{I \in \mathcal{C}_n} (\Delta_I \psi)^+ + \sum_{I \in \mathcal{C}_n} (\Delta_I \psi)^- \right) = \mathrm{var}^+_{[a,b]}(\psi) + \mathrm{var}^-_{[a.b]}(\psi).$$

□

Given a function ψ of bounded variation on $[a, b]$, define $V_\psi(x) = \mathrm{var}_{[a,x]}(\psi)$ and $V^{\pm}_\psi(x) = \mathrm{var}^{\pm}_{[a,x]}(\psi)$ for $x \in [a, b]$. Then V_ψ, V^+_ψ, and V^-_ψ are all non-decreasing functions that vanish at a, and, by Lemma 3.5.3, $\psi(x) = \psi(a) + V^+_\psi(x) - V^-_\psi(x)$ and $V_\psi(x) = V^+_\psi(x) + V^-_\psi(x)$ for $x \in [a, b]$.

Theorem 3.5.4 *Let ψ be a function of bounded variation on $[a, b]$, and refer to the preceding. If $\varphi : [a, b] \longrightarrow \mathbb{R}$ is a bounded function, then φ is Riemann–Stieltjes* *integrable with respect to* V_ψ *if and only if it is Riemann–Stieltjes integrable with respect to both* V_ψ^+ *and* V_ψ^-, *in which case*

$$\int_a^b \varphi(x)\,dV_\psi(x) = \int_a^b \varphi(x)\,dV_\psi^+(x) + \int_a^b \varphi(x)\,dV_\psi^-(x).$$

Moreover, if φ *is Riemann–Stieltjes integrable with respect to* V_ψ, *then it is Riemann–Stieltjes integrable with respect to* ψ,

$$\int_a^b \varphi(x)\,d\psi(x) = \int_a^b \varphi(x)\,dV_\psi^+(x) - \int_a^b \varphi(x)\,dV_\psi^-(x),$$

and

$$\left|\int_a^b \varphi(x)\,d\psi(x)\right| \le \int_a^b |\varphi(x)|\,dV_\psi(x) \le \|\varphi\|_{[a,b]}\mathrm{var}_{[a,b]}(\psi).$$

Proof Since $V_\psi = V_\psi^+ + V_\psi^-$, it is clear that

$$\sum_{\substack{I\in\mathcal{C}\\ \sup_I \varphi - \inf_I \varphi > \epsilon}} \Delta V_\psi < \epsilon \iff \sum_{\substack{I\in\mathcal{C}\\ \sup_I \varphi - \inf_I \varphi > \epsilon}} \Delta V_\psi^+ + \sum_{\substack{I\in\mathcal{C}\\ \sup_I \varphi - \inf_I \varphi > \epsilon}} \Delta V_\psi^- < \epsilon,$$

and therefore, by Lemma 3.5.2, φ is Riemann–Stieltjes integrable with respect to V_ψ if and only if it is with respect to V_ψ^+ and V_ψ^-. Furthermore, because

$$\mathfrak{R}(\varphi|V_\psi;\mathcal{C},\Xi) = \mathfrak{R}(\varphi|V_\psi^+;\mathcal{C},\Xi) + \mathfrak{R}(\varphi|V_\psi^-;\mathcal{C},\Xi)$$

and

$$\mathfrak{R}(\varphi|\psi;\mathcal{C},\Xi) = \mathfrak{R}(\varphi|V_\psi^+;\mathcal{C},\Xi) - \mathfrak{R}(\varphi|V_\psi^-;\mathcal{C},\Xi),$$

the other assertions follow easily. □

Finally, there is an important case in which Riemann–Stieltjes integrals reduce to Riemann integrals. Namely, if ψ is continuous on $[a,b]$ and continuously differentiable on (a,b), then, by (1.8.1),

$$\sum_{I\in\mathcal{C}} |\Delta_I\psi| \le \|\psi'\|_{(a.b)}(b-a),$$

and so ψ will have bounded variation if ψ' is bounded. Furthermore, if ψ' is bounded and $\varphi : [a,b] \longrightarrow \mathbb{R}$ is a bounded function which is Riemann integrable on $[a,b]$, then φ is Riemann–Stieltjes integrable with respect to ψ and

$$\int_a^b \varphi(x)\,d\psi(x) = \int_a^b \varphi(x)\psi'(x)\,dx. \tag{3.5.3}$$

To prove this, note that if $I \in \mathcal{C}$, then one can apply (1.8.1) to find an $\eta(I) \in I$ such that $\Delta_I\psi = \psi'\big(\eta(I)\big)|I|$. Thus

$$\mathfrak{R}(\varphi|\psi;\mathcal{C},\Xi) = \sum_{I\in\mathcal{C}} \varphi\big(\eta(I)\big)\psi'\big(\eta(I)\big)|I| + \sum_{I\in\mathcal{C}}\Big(\varphi\big(\Xi(I)\big) - \varphi\big(\eta(I)\big)\Big)\psi'\big(\eta(I)\big)|I|$$

and

$$\left|\sum_{I\in\mathcal{C}}\Big(\varphi\big(\Xi(I)\big) - \varphi\big(\eta(I)\big)\Big)\psi'\big(\eta(I)\big)|I|\right| \le \|\psi'\|_{(a,b)}\big(\mathcal{U}(\varphi;\mathcal{C}) - \mathcal{L}(\varphi;\mathcal{C})\big),$$

which, since φ is Riemann integrable, tends to 0 as $\|\mathcal{C}\| \to 0$. Hence, since $\varphi\psi'$, as the product of two Riemann integrable functions, is Riemann integrable, we see that $\mathfrak{R}(\varphi|\psi;\mathcal{C},\Xi) \longrightarrow \int_a^b \varphi(x)\psi'(x)\,dx$ as $\|\mathcal{C}\| \to 0$. The content of (3.5.3) is often abbreviated by the equation $d\psi(x) = \psi'(x)\,dx$. Notice that when φ and ψ are both continuously differentiable on (a,b), then (3.5.2) is the integration by parts formula.

3.6 Exercises

Exercise 3.1 Most integrals defy computation. Here are a few that don't. In each case, compute the following integrals.

(i) $\int_{[1,\infty)} \frac{1}{x^2}\,dx$	(ii) $\int_{[0,\infty)} e^{-x}\,dx$
(iii) $\int_a^b \sin x\,dx$	(iv) $\int_a^b \cos x\,dx$
(v) $\int_0^1 \frac{1}{\sqrt{1-x^2}}\,dx$	(vi) $\int_{[0,\infty)} \frac{1}{1+x^2}\,dx$
(vii) $\int_0^{\frac{\pi}{2}} x^2\sin x\,dx$	(viii) $\int_0^1 \frac{x}{x^4+1}\,dx$

Exercise 3.2 Here are some integrals that play a role in Fourier analysis. In evaluating them, it may be helpful to make use of (1.5.1). Compute

(i) $\int_0^{2\pi} \sin(mx)\cos(nx)\,dx$ (ii) $\int_0^{2\pi} \sin^2(mx)\,dx$ (iii) $\int_0^{2\pi} \cos^2(mx)\,dx$,

for $m,\ n \in \mathbb{N}$.

Exercise 3.3 For $t > 0$, define *Euler's Gamma function* $\Gamma(t)$ for $t > 0$ by

$$\Gamma(t) = \int_{(0,\infty)} x^{t-1}e^{-x}\,dx.$$

Show that $\Gamma(t+1) = t\Gamma(t)$, and conclude that $\Gamma(n+1) = n!$ for $n \geq 1$. See (5.4.4) for the evaluation of $\Gamma\big(\frac{1}{2}\big)$.

Exercise 3.4 Find indefinite integrals for the following functions:

(i) x^α for $\alpha \in \mathbb{R}$ & $x \in (0,\infty)$ (ii) $\log x$
(iii) $\frac{1}{x \log x}$ for $x \in (0,\infty) \setminus \{1\}$ (iv) $\frac{(\log x)^n}{x}$ for $n \in \mathbb{Z}^+$ & $x \in (0,\infty)$.

Exercise 3.5 Let $\alpha, \beta \in \mathbb{R}$, and assume that $(\alpha a) \vee (\alpha b) \vee (\beta a) \vee (\beta b) < 1$. Compute $\int_a^b \frac{1}{(1-\alpha x)(1-\beta x)}\,dx$. When $\alpha = \beta = 0$, there is nothing to do, and when $\alpha = \beta \neq 0$, it is obvious that $\big(\alpha(1-\alpha x)\big)^{-1}$ is an indefinite integral. When $\alpha \neq \beta$, one can use the method of *partial fractions* and write

$$\frac{1}{(1-\alpha x)(1-\beta x)} = \frac{1}{\alpha-\beta}\left(\frac{\alpha}{1-\alpha x} - \frac{\beta}{1-\beta x}\right).$$

See Theorem 6.3.2 for a general formulation of this procedure.

Exercise 3.6 Suppose that $f : (a,b) \longrightarrow [0,\infty)$ has continuous derivatives of all orders. Then f is said to be *absolutely monotone* if it and all its derivatives are non-negative. If f is absolutely monotone, show that for each $c \in (a,b)$

$$f(x) = \sum_{m=0}^{\infty} \frac{f^{(m)}(c)}{m!}(x-c)^m \quad \text{for } x \in [c,b),$$

an observation due to S. Bernstein. In doing this problem, reduce to the case when $c = 0 \in (a,b)$, and, using (3.2.2), observe that $f(y)$ dominates

$$\frac{y^n}{(n-1)!}\int_0^1 (1-t)^{n-1} f^{(n)}(ty)\,dt \geq \left(\frac{y}{x}\right)^n \frac{x^n}{(n-1)!}\int_0^1 (1-t)^{n-1} f^{(n)}(tx)\,dt$$

for $n \geq 1$ and $0 < x < y < b$.

Exercise 3.7 For $\alpha \in \mathbb{R} \setminus \{0\}$, show that

$$\int_a^b e^{i\alpha t}\,dt = \frac{e^{i\alpha b} - e^{i\alpha a}}{i\alpha}.$$

Next, write $\cos t = \frac{e^{it}+e^{-it}}{2}$, and apply the preceding and the binomial formula to show that

$$\int_0^{2\pi} \cos^n t\,dt = \begin{cases} 0 & \text{if } n \in \mathbb{N} \text{ is odd} \\ 2^{-n+1}\pi\binom{n}{\frac{n}{2}} & \text{if } n \in \mathbb{N} \text{ is even.} \end{cases}$$

By combining this with (3.2.4), show that $\lim_{n\to\infty} n^{\frac{1}{2}} \int_0^{2\pi} \cos^{2n} t\,dt = 2\pi^{\frac{1}{2}}$.

Exercise 3.8 A more conventional way to introduce the logarithm function is to define it by

$$(*) \qquad \log y = \int_1^y \frac{1}{t}\,dt \quad \text{for } y \in (0,\infty).$$

The purpose of this exercise is to show, without knowing our earlier definition, that this definition works.

(**i**) Suppose that $\ell : (0,\infty) \longrightarrow \mathbb{R}$ is a continuous function with the properties that $\ell(xy) = \ell(x) + \ell(y)$ for all $x,\ y \in (0,\infty)$ and $\ell(a) > 0$ for some $a > 1$. By applying Exercise 1.10 to the function $f(x) = \ell(a^x)$, show that $\ell\big(a^x\big) = x\ell(a)$.

(**ii**) Referring to (**i**), show that ℓ is strictly increasing, tends to ∞ as $y \to \infty$ and to $-\infty$ as $y \searrow 0$. Conclude that there is a unique $b \in (1,\infty)$ such that $\ell(b) = 1$.

(**iii**) Continuing (**i**) and (**ii**), and again using Exercise 1.10, conclude that $\ell(b^x) = x$ for $x \in \mathbb{R}$ and $b^{\ell(y)} = y$ for $y \in (0,\infty)$. That is, ℓ is the logarithm function with base b

(**iv**) Show that the function log given by $(*)$ satisfies the conditions in (**i**) and therefore that there exists a unique $e \in (1,\infty)$ for which it is the logarithm function with base e.

Exercise 3.9 Show that

$$\lim_{x\to\infty} \frac{\log x}{x} \int_e^x \frac{1}{\log t}\,dt = 1.$$

Exercise 3.10 Let $f : [0,1] \longrightarrow \mathbb{C}$ be a continuous function whose periodic extension is continuously differentiable. Show that

$$\int_0^1 \left| f(x) - \int_0^1 f(y)\,dy \right|^2 dx = \int_0^1 |f(x)|^2\,dx - \left| \int_0^1 f(y)\,dy \right|^2 = \sum_{m\neq 0} |\hat{f}_m|^2$$
$$\leq (2\pi)^{-2} \sum_{m\neq 0} (2\pi m)^2 |\hat{f}_m|^2 = (2\pi)^{-2} \int_0^1 |f'(x)|^2\,dx.$$

As a consequence, one has the *Poincaré inequality*

$$\int_0^1 \left| f(x) - \int_0^1 f(y)\,dy \right|^2 dx \leq (2\pi)^{-2} \int_0^1 |f'(x)|^2\,dx$$

for any function whose periodic extension is continuously differentiable.

Exercise 3.11 Let $f : [a,b] \longrightarrow \mathbb{C}$ be a continuous function, and set $L = b - a$. If $\delta \in \big(0, \frac{L}{2}\big)$, show that, as $r \nearrow 1$,

$$\frac{1}{L}\sum_{m=-\infty}^{\infty} r^{|m|}\left(\int_a^b f(y)\mathfrak{e}_{-m}\left(\tfrac{y}{L}\right)dy\right)\mathfrak{e}_m\left(\tfrac{x}{L}\right)$$

converges to $f(x)$ uniformly for $x \in \left[a+\delta, b-\delta\right]$ and that the convergence is uniform for $x \in [a,b]$ if $f(b) = f(a)$. Perhaps the easiest way to do this is to consider the function $g(x) = f(a+Lx)$ and apply Theorem 3.4.1 to it. Next, assume that $f(a) = f(b)$ and that f has a bounded, continuous derivative on (a,b). Show that

$$f(x) = \frac{1}{L}\sum_{m=-\infty}^{\infty}\left(\int_a^b f(y)\mathfrak{e}_{-m}\left(\tfrac{y}{L}\right)dy\right)\mathfrak{e}_m\left(\tfrac{x}{L}\right),$$

where the convergence is absolute and uniform on $[a,b]$.

Exercise 3.12 Let $f : [0,1] \longrightarrow \mathbb{C}$ be a continuous function, and show that, as $r \nearrow 1$,

$$f_r(x) \equiv 2\sum_{m=1}^{\infty} r^m\left(\int_0^1 f(y)\sin(m\pi y)\,dy\right)\sin \pi x$$

converges uniformly to f on $[\delta, 1-\delta]$ for each $\delta \in \left(0, \frac{1}{2}\right)$. One way to do this is to define $g : [-1,1] \longrightarrow \mathbb{C}$ so that $g = f$ on $[0,1]$ and $g(x) = -f(-x)$ when $x \in [-1,0)$, and observe that

$$\int_{-1}^1 g(y)\mathfrak{e}_{-m}\left(\tfrac{y}{2}\right)dy = -2i\int_0^1 f(y)\sin(m\pi y)\,dy.$$

If $f(0) = 0$ and therefore g is continuous, one need only apply Exercise 3.11 to g to see that $f_r \longrightarrow f$ uniformly on $[0, 1-\delta]$ for each $\delta \in \left(0, \frac{1}{2}\right)$. When $f(0) \neq$ and therefore g is discontinuous at 0, after examining the proof of Theorem 3.4.1, one can show that $f_r \longrightarrow f$ on $[\delta, 1-\delta]$ because g is uniformly continuous there.

Next show that if $f : [0,1] \longrightarrow \mathbb{C}$ is continuous, then

$$\int_0^1 |f(x)|^2\,dx = 2\sum_{m=1}^{\infty}\left|\int_0^1 f(x)\sin(m\pi x)\,dx\right|^2.$$

Finally, assuming that $f(0) = 0 = f(1)$ and that f has a bounded, continuous derivative on $(0,1)$, show that

$$f(x) = 2\sum_{m=1}^{\infty}\left(\int_0^1 f(y)\sin \pi y\,dy\right)\sin \pi x,$$

where the convergence of the series is absolute and uniform on $[0,1]$.

Exercise 3.13 Fourier series provide an alternative and more elegant approach to proving estimates like the one in (3.3.5). To see this, suppose that $f : [0, 1] \longrightarrow \mathbb{C}$ is a function whose periodic extension is $\ell \geq 1$ times continuously differentiable. Then, as we have shown,

$$f(x) = \int_0^1 f(y)\,dy + \sum_{m\neq 0} \frac{\widehat{(f^{(\ell)})}_m}{(i2\pi m)^\ell} \mathfrak{e}_m(x),$$

where the series converges uniformly and absolutely. After showing that

$$\frac{1}{n}\sum_{k=1}^{n} \mathfrak{e}_m\left(\tfrac{k}{n}\right) = \begin{cases} 1 & \text{if } m \text{ is divisible by } n \\ 0 & \text{otherwise,} \end{cases}$$

conclude that

$$\mathcal{R}_n(f) - \int_0^1 f(y)\,dy = \sum_{m\neq 0} \frac{\widehat{(f^{(\ell)})}_{mn}}{(i2\pi mn)^\ell},$$

and from this show that

$$\left|\mathcal{R}_n(f) - \int_0^1 f(y)\,dy\right| \leq \frac{2\|f^{(\ell)}\|_{[0,1]}\zeta(\ell)}{(2\pi n)^\ell}.$$

Finally, use Schwarz's inequality and (3.4.5) to derive the estimate

$$\left|\mathcal{R}_n(f) - \int_0^1 f(y)\,dy\right| \leq \frac{\sqrt{2\zeta(2\ell)}}{(2\pi n)^\ell}\sqrt{\int_0^1 |f(x)|^2\,dx}.$$

Exercise 3.14 Think of the unit circle $\mathbb{S}^1(0, 1)$ as a subset of $\mathbb{C}$, and let $\varphi : \mathbb{S}^1(0, 1) \longrightarrow \mathbb{R}$ be a continuous function. The goal of this exercise is to show that there exists an analytic function f on $D(0, 1)$ such that $\lim_{z\to\zeta} \mathfrak{R}\big(f(z)\big) = \varphi(\zeta)$ for all $\zeta \in \mathbb{S}^1(0, 1)$.

(**i**) Set

$$a_m = \int_0^1 \varphi\big(e^{i2\pi\theta}\big)\mathfrak{e}_{-m}(\theta)\,d\theta \quad \text{for } m \in \mathbb{Z},$$

and show that $\overline{a_m} = a_{-m}$. Next define the function u on $D(0, 1)$ by

$$u\big(re^{i2\pi}\big) = \sum_{m=-\infty}^{\infty} r^{|m|} a_m \mathfrak{e}_m(\theta) \quad \text{for } r \in [0, 1) \text{ and } \theta \in [0, 1).$$

Show that u is a continuous, $\mathbb{R}$-valued function and, using Theorem 3.4.1, that $\lim_{z\to\zeta} u(z) = \varphi(\zeta)$ for $\zeta \in \mathbb{S}^1(0, 1)$.

(**ii**) Define v on $D(0,1)$ by

$$v\big(re^{i2\pi\theta}\big) = -i\sum_{m=1}^{\infty} r^m\big(a_m\mathfrak{e}_m(\theta) - a_{-m}\mathfrak{e}_{-m}(\theta)\big).$$

and show that v is a continuous, $\mathbb{R}$-valued function. Next, set $f = u + iv$, and show that

$$f(z) = a_0 + 2\sum_{m=1}^{\infty} a_m z^m \quad \text{for } z \in D(0,1).$$

In particular, conclude that f is analytic and that $\lim_{z\to\zeta}\Re\big(f(z)\big) = \varphi(\zeta)$ for $\zeta \in \mathbb{S}^1(0,1)$.

(**iii**) Assume that $\sum_{m=-\infty}^{\infty}|a_m| < \infty$, and define $H\varphi$ on $\mathbb{S}^1(0,1)$ by

$$H\varphi\big(e^{i2\pi\theta}\big) = -i\sum_{m=1}^{\infty}\big(a_m\mathfrak{e}_m(\theta) - a_{-m}\mathfrak{e}_{-m}(\theta)\big).$$

If f is the function in (**ii**), show that $\lim_{z\to\zeta}\Im\big(f(z)\big) = H\varphi(\zeta)$ for $\zeta \in \mathbb{S}^1(0,1)$. The function $H\varphi$ is called the *Hilbert transform* of φ, and it plays an important role in harmonic analysis.

Exercise 3.15 Let $\{x_n : n \ge 1\}$ be a sequence of distinct elements of $(a,b]$, and set $S(x) = \{n \ge 1 : x_n \le x\}$ for $x \in [a,b]$. Given a sequence $\{c_n : n \ge 1\} \subseteq \mathbb{R}$ for which $\sum_{n=1}^{\infty}|c_n| < \infty$, define $\psi : [a,b] \longrightarrow \mathbb{R}$ so that $\psi(x) = \sum_{n\in S(x)} c_n$. Show that ψ has bounded variation and that $\|\psi\|_{\text{var}} = \sum_{n=1}^{\infty}|c_n|$. In addition, show that if $\varphi : [a,b] \longrightarrow \mathbb{R}$ is continuous, then

$$\int_a^b \varphi(x)\,d\psi(x) = \sum_{n=1}^{\infty}\varphi(x_n)c_n.$$

Exercise 3.16 Suppose that $F : [a,b] \longrightarrow \mathbb{R}$ is a function of bounded variation.

(**i**) Show that, for each $\epsilon > 0$,

$$\varlimsup_{h\searrow 0}\big|F(x+h) - F(x)\big| \vee |F(x-h) - F(x)| \ge \epsilon$$

for at most a finite number of $x \in (a,b)$, and use this to show that F is Riemann integrable on $[a,b]$.

(**ii**) Assume that $\varphi : [a,b] \longrightarrow \mathbb{R}$ is a continuous function that has a bounded, continuous derivative on (a,b). Prove the *integration by parts* formula

$$\int_a^b \varphi(t)\,dF(t) = \varphi(b)F(b) - \varphi(a)F(a) - \int_a^b \varphi'(t)F(t)\,dt.$$

(**iii**) Let $\psi : [0,\infty) \longrightarrow [0,\infty)$ be a function whose restriction to $[0,T]$ has bounded variation for each $T > 0$. Assuming that $\psi(0) = 0$ and that $\sup_{t\geq 1} t^{-\alpha}|\psi(t)| < \infty$ for some $\alpha \geq 0$, use (**ii**) to show that

$$L(\lambda) \equiv \int_{[0,\infty)} e^{-\lambda t}\, d\psi(t) = \lim_{T\to\infty} \int_0^T e^{-\lambda t}\, d\psi(t) = \lambda \int_{[0,\infty)} e^{-\lambda t} \psi(t)\, dt$$

for all $\lambda > 0$. The function $\lambda \rightsquigarrow L(\lambda)$ is called the *Laplace transform* of ψ.

(**iv**) Refer to (**iii**), and assume that $a = \lim_{t\to\infty} t^{-\alpha}\psi(t) \in \mathbb{R}$ exists. Show that, for each $T > 0$, $\lambda^\alpha L(\lambda)$ equals

$$\lambda^{1+\alpha} \int_0^T e^{-\lambda t}\psi(t)\, dt + \lambda^{1+\alpha} \int_{[T,\infty)} t^\alpha e^{-\lambda t}\big(t^{-\alpha}\psi(t) - a\big) dt + a \int_{[\lambda T,\infty)}^{\infty} t^\alpha e^{-t}\, dt,$$

and use this to conclude that (cf. Exercise 3.3)

$$(*) \qquad a = \lim_{\lambda \searrow 0} \frac{\lambda^\alpha L(\lambda)}{\Gamma(1+\alpha)}.$$

This equation is an example of the general principle that the behavior of ψ near infinity reflects the behavior of its Laplace transform at 0.

(**v**) The equation in $(*)$ is an integral version of the *Abel limit* procedure in 1.10.1, and it generalizes 1.10.1. To see this, let $\{c_n : n \geq 1\} \subseteq \mathbb{R}$ be given, and define

$$\psi(t) = \sum_{1\leq n\leq t} c_n \quad \text{for } t \in [0,\infty).$$

Assuming that $\sup_{n\geq 1} n^{-\alpha}|\psi(n)| < \infty$, use Exercise 3.15 to show that

$$\int_{[0,\infty)} e^{-\lambda t}\, d\psi(t) = \sum_{n=1}^{\infty} e^{-\lambda n} c_n \quad \text{for } \lambda > 0.$$

Next, assume that $a = \lim_{n\to\infty} n^{-\alpha}\psi(n) \in \mathbb{R}$ exists, and use $(*)$ to conclude that

$$a = \lim_{\lambda \searrow 0} \frac{\lambda^\alpha}{\Gamma(1+\alpha)} \sum_{n=1}^{\infty} e^{-\lambda n} c_n.$$

When $\alpha = 0$, this is equivalent to 1.10.1.

Chapter 4
Higher Dimensions

Many of the ideas that we developed for $\mathbb{R}$ and $\mathbb{C}$ apply equally well to $\mathbb{R}^N$ when $N \geq 3$. However, when $N \geq 3$, there is no multiplicative structure that has the properties that ordinary multiplication has in $\mathbb{R}$ or that the multiplication that we introduced in $\mathbb{R}^2$ has there. Nonetheless, there is a natural linear structure. That is, if $\mathbf{x} = (x_1, \dots, x_N)$ and $\mathbf{y} = (y_1, \dots, y_N)$ are elements of $\mathbb{R}^N$, then for any $\alpha, \beta \in \mathbb{R}$, the operation $\alpha\mathbf{x} + \beta\mathbf{y} \equiv (\alpha x_1 + \beta y_1, \dots, \alpha y_N + \beta y_N)$ turns $\mathbb{R}^N$ into a linear space. Further, the geometric interpretation of this operation is the same as it was for $\mathbb{R}^2$. To be precise, think of $\mathbf{x}$ as a vector based at the origin and pointing toward $(x_1, \dots, x_N) \in \mathbb{R}^N$. Then, $-\mathbf{x}$ is the vector obtained by reversing the direction of $\mathbf{x}$, and, when $\alpha \geq 0$, $\alpha\mathbf{x}$ is obtained from $\mathbf{x}$ by rescaling its length by the factor α. When $\alpha < 0$, $\alpha\mathbf{x} = |\alpha|(-\mathbf{x})$ is the vector obtained by first reversing the direction of $\mathbf{x}$ and then rescaling its length by the factor of $|\alpha|$. Finally, if the end of the vector corresponding to $\mathbf{y}$ is moved to the point of the vector corresponding to $\mathbf{x}$, then the point of the $\mathbf{y}$-vector will be at $\mathbf{x} + \mathbf{y}$.

4.1 Topology in $\mathbb{R}^N$

In addition to its linear structure, $\mathbb{R}^N$ has a natural notion of length. Namely, given $\mathbf{x} = (x_1, \dots, x_N)$, define its *Euclidean length* $|\mathbf{x}| = \sqrt{x_1^2 + \cdots + x_n^2}$. The analysis of this quantity is simplified by the introduction of what is called the *inner product* $(\mathbf{x}, \mathbf{y})_{\mathbb{R}^N} \equiv \sum_{j=1}^N x_j y_j$. Obviously, $|\mathbf{x}|^2 = (\mathbf{x}, \mathbf{x})_{\mathbb{R}^N}$. Further, by Schwarz's inequality (cf. Exercise 2.3), it is clear that

$$|(\mathbf{x}, \mathbf{y})_{\mathbb{R}^N}| \leq \sum_{j=1}^N |x_j||y_j| \leq |\mathbf{x}||\mathbf{y}|,$$

D.W. Stroock, *A Concise Introduction to Analysis*,
DOI 10.1007/978-3-319-24469-3_4

which means that $|(\mathbf{x},\mathbf{y})_{\mathbb{R}^N}| \le |\mathbf{x}||\mathbf{y}|$, an inequality that is also called *Schwarz's inequality* or sometimes the *Cauchy–Schwarz inequality*. Further, equality holds in Schwarz's inequality if and only if there is a $t \in \mathbb{R}$ such that either $\mathbf{y} = t\mathbf{x}$ or $\mathbf{x} = t\mathbf{y}$. If $\mathbf{x} = \mathbf{0}$, this is completely trivial. Thus, assume that $\mathbf{x} \ne \mathbf{0}$, and notice that, for any $t \in \mathbb{R}$, $\mathbf{y} = t\mathbf{x}$ if and only if

$$0 = |\mathbf{y} - t\mathbf{x}|^2 = t^2|\mathbf{x}|^2 - 2t(\mathbf{x},\mathbf{y})_{\mathbb{R}^N} + |\mathbf{y}|^2.$$

By the quadratic formula, such a t must equal

$$\frac{(\mathbf{x},\mathbf{y})_{\mathbb{R}^N} \pm \sqrt{(\mathbf{x},\mathbf{y})^2_{\mathbb{R}^N} - |\mathbf{x}|^2|\mathbf{y}|^2}}{|\mathbf{x}|^2},$$

which, since $t \in \mathbb{R}$, is possible if and only if $(\mathbf{x},\mathbf{y})^2_{\mathbb{R}^N} - |\mathbf{x}|^2|\mathbf{y}|^2 \ge 0$. Since $(\mathbf{x},\mathbf{y})^2_{\mathbb{R}^N} \le |\mathbf{x}|^2\mathbf{y}|^2$, if follows that t exists if and only if $(\mathbf{x},\mathbf{y})^2_{\mathbb{R}^N} = |\mathbf{x}|^2|\mathbf{y}|^2$.

Knowing Schwarz's inequality, we see that

$$|\mathbf{x}+\mathbf{y}|^2 = |\mathbf{x}|^2 + 2(\mathbf{x},\mathbf{y})_{\mathbb{R}^N} + |\mathbf{y}|^2 \le |\mathbf{x}|^2 + 2|\mathbf{x}||\mathbf{y}| + |\mathbf{y}|^2 = (|\mathbf{x}| + |\mathbf{y}|)^2,$$

and therefore $|\mathbf{x}+\mathbf{y}| \le |\mathbf{x}| + |\mathbf{y}|$, which is known as the *triangle inequality* since, when thought about in terms of vectors, it says that the length of the third side of a triangle is less than or equal to the sum of the lengths of the other two sides.

Just as we did for $\mathbb{R}$ and $\mathbb{R}^2$, we will say the sequence $\{\mathbf{x}_n : n \ge 1\} \subseteq \mathbb{R}^N$ converges to $\mathbf{x} \in \mathbb{R}^N$ if for each $\epsilon > 0$ there exists an n_ϵ such that $|\mathbf{x}_n - \mathbf{x}| < \epsilon$ for all $n \ge n_\epsilon$, in which case we write $\mathbf{x}_n \longrightarrow \mathbf{x}$ or $\mathbf{x} = \lim_{n\to\infty} \mathbf{x}_n$. Writing $\mathbf{x}_n = (x_{1,n}, \dots, x_{N,n})$ and $\mathbf{x} = (x_1, \dots, x_N)$, it is easy to check that $\mathbf{x}_n \longrightarrow \mathbf{x}$ if and only if $|x_{j,n} - x_j| \longrightarrow 0$ for each $1 \le j \le N$. Thus, since $\lim_{m\to\infty} \sup_{n\ge m} |\mathbf{x}_n - \mathbf{x}_m| = 0$ if and only if $\lim_{m\to\infty} \sup_{n\ge m} |x_{j,n} - x_{j,m}| = 0$ for each $1 \le j \le N$, it follows that Cauchy's criterion holds in this context. That is, $\mathbb{R}^N$ is complete.

Given $\mathbf{x}$ and $r > 0$, the *open ball* $B(\mathbf{x}, r)$ of radius r is the set of $\mathbf{y} \in \mathbb{R}^N$ such that $|\mathbf{y} - \mathbf{x}| < r$. A set $G \subseteq \mathbb{R}^N$ is said to be *open* if, for each $\mathbf{x} \in G$ there is an $r > 0$ such that $B(\mathbf{x}, r) \subseteq G$, and a subset F is said to be *closed* if its complement is open. In particular, $\mathbb{R}^N$ and $\emptyset$ are both open and closed. Further, the *interior* $\mathrm{int}(S)$ and *closure* $\bar{S}$ of an $S \subseteq \mathbb{R}^N$ are, respectively, the largest open set contained in S and the smallest closed set containing it. By the same reasoning as was used to prove Lemma 1.3.1, $\mathbf{x} \in \mathrm{int}(S)$ if and only if $B(\mathbf{x}, r) \subseteq S$ for some $r > 0$ and $\mathbf{x} \in \bar{S}$ if and only if there is a $\{\mathbf{x}_n : n \ge 1\} \subseteq S$ such that $\mathbf{x}_n \longrightarrow \mathbf{x}$.

A set $K \subseteq \mathbb{R}^N$ is said to be *bounded* if $\sup_{\mathbf{x}\in K} |\mathbf{x}| < \infty$, and it is said to be *compact* if every sequence $\{\mathbf{x}_n : n \ge 1\} \subseteq K$ admits a subsequence $\{\mathbf{x}_{n_k} : k \ge 1\}$ that converges to some point $\mathbf{x} \in K$.

The following lemma is sometimes called the *Heine–Borel theorem*.

Lemma 4.1.1 *If $\{\mathbf{x}_n : n \geq 1\}$ is a bounded sequence in $\mathbb{R}^N$, then it admits a convergent subsequence $\{\mathbf{x}_{n_k} : k \geq 1\}$. In particular, $K \subseteq \mathbb{R}^N$ is compact if and only if K is closed and bounded. (See Exercise 4.9 for more information.)*

Proof Suppose $\mathbf{x}_n = (x_{1,n}, \ldots, x_{N,n})$ for each $n \geq 1$. Then, for each $1 \leq j \leq N$, $\{x_{j,n} : n \geq 1\}$ is a bounded sequence in $\mathbb{R}$, and so, by Theorem 1.3.3, there exists a subsequence $\{x_{1,n_k^{(1)}} : k \geq 1\}$ of $\{x_{1,n} : n \geq 1\}$ that converges to an $x_1 \in \mathbb{R}$. Next, again by Theorem 1.3.3, there is a subsequence $\{x_{2,n_k^{(2)}} : k \geq 1\}$ of the sequence $\{x_{1,n_k^{(1)}} : k \geq 1\}$ that converges to an $x_2 \in \mathbb{R}$. Proceeding in this way, we can produce subsequences $\{\mathbf{x}_{n_k^{(j)}} : k \geq 1\}$ for $1 \leq j \leq N$ such that, for each $1 \leq j < N$, $\{\mathbf{x}_{n_k^{(j+1)}} : k \geq 1\}$ is a subsequence of $\{\mathbf{x}_{n_k^{(j)}} : k \geq 1\}$ and $x_{j,n_k^{(j)}} \longrightarrow x_j \in \mathbb{R}$ for each $1 \leq j \leq N$. Hence, $\{\mathbf{x}_{n_k^{(N)}} : k \geq 1\}$ converges to $\mathbf{x} = (x_1, \ldots, x_N)$.

Now suppose that $K \subseteq \mathbb{R}^N$, and let $\{\mathbf{x}_n : n \geq 1\} \subseteq K$. If K is closed and bounded, then $\{\mathbf{x}_n : n \geq 1\}$ is bounded and therefore, by the preceding, admits a convergent subsequence whose limit is necessarily is in K since K is closed. Conversely, if K is compact and $\mathbf{x}_n \longrightarrow \mathbf{x}$, then $\mathbf{x}$ must be in K since there is a subsequence of $\{\mathbf{x}_n : n \geq 1\}$ that converges to a point in K. In addition, if K were unbounded, then there would be a sequence $\{\mathbf{x}_n : n \geq 1\} \subseteq K$ such that $|\mathbf{x}_n| \geq n$ for all $n \geq 1$. Further, because K is compact, $\{\mathbf{x}_n : n \geq 1\}$ could be chosen so that it converges to a point $\mathbf{x} \in K$. But then $|\mathbf{x}| = \infty$, and so no such sequence can exist. □

If $\emptyset \neq S \subseteq \mathbb{R}^N$ and $f : S \longrightarrow \mathbb{C}$, then f is said to be *continuous* at a point $x \in S$ if for each $\epsilon > 0$ there is a $\delta > 0$ with the property that $|f(\mathbf{y}) - f(\mathbf{x})| < \epsilon$ whenever $\mathbf{y} \in S \cap B(\mathbf{x}, \delta)$. Just as in Lemma 1.3.5, f is continuous at $\mathbf{x} \in S$ if and only if $f(\mathbf{x}_n) \longrightarrow f(\mathbf{x})$ whenever $\{\mathbf{x}_n : n \geq 1\} \subseteq S$ converges to $\mathbf{x}$, and, when S is open, f is continuous on S if and only if $f^{-1}(G)$ is open for all open $G \subseteq \mathbb{C}$. Also, using Lemma 4.1.1 and arguing in the same way as we did in Theorem 1.4.1, one can show that if f is a continuous function on a compact subset K, then it is bounded (i.e., $\sup_{\mathbf{x} \in S} |f(\mathbf{x})| < \infty$) and uniformly continuous in the sense that, for each $\epsilon > 0$, there is a $\delta > 0$ such that $|f(\mathbf{y}) - f(\mathbf{x})| < \epsilon$ whenever $\mathbf{x} \in K$ and $\mathbf{y} \in K \cap B(\mathbf{x}, \delta)$. Further, if f is $\mathbb{R}$-valued, it achieves both its maximum and a minimum values, and the reasoning in Lemma 1.4.4 applies equally well here and shows that if $\{f_n : n \geq 1\}$ is a sequence of continuous functions on $S \subseteq \mathbb{R}^N$ and if $\{f_n : n \geq 1\}$ converges uniformly on S to a function f, then f is also continuous. Finally, if $\mathbf{F} = (F_1, \ldots, F_M) : S \longrightarrow \mathbb{R}^M$, then we say that $\mathbf{F}$ is continuous at $\mathbf{x}$ if for each $\epsilon > 0$ there is a $\delta > 0$ such that $|\mathbf{F}(\mathbf{y}) - \mathbf{F}(\mathbf{x})| < \epsilon$ when $\mathbf{y} \in S \cap B(\mathbf{x}, \delta)$. It is easy to check that $\mathbf{F}$ is continuous at $\mathbf{x}$ if and only if F_j is for each $1 \leq j \leq M$.

Say that $S \subseteq \mathbb{R}^N$ is *path connected* (cf. Exercise 4.5) if for all $\mathbf{x}_0$ and $\mathbf{x}_1$ in S there is a continuous path $\gamma : [0, 1] \longrightarrow S$ such that $\mathbf{x}_0 = \gamma(0)$ and $\mathbf{x}_1 = \gamma(1)$.

Lemma 4.1.2 *Assume that S is path connected and that $f : S \longrightarrow \mathbb{R}$ is continuous. For all $\mathbf{x}_0, \mathbf{x}_1 \in S$ and $y \in \big[f(\mathbf{x}_0) \wedge f(\mathbf{x}_1), f(\mathbf{x}_0) \vee f(\mathbf{x}_1)\big]$, there exists an $\mathbf{x} \in S$ such that $f(\mathbf{x}) = y$. In particular, if K is a path connected, compact subset of $\mathbb{R}^N$ and $f : K \longrightarrow \mathbb{R}$ is continuous, then for all $y \in [m, M]$ there is an $\mathbf{x} \in K$ such that $f(\mathbf{x}) = y$ where $m = \min\{f(\mathbf{x}) : \mathbf{x} \in K\}$ and $M = \max\{f(\mathbf{x}) : \mathbf{x} \in K\}$.*

Proof To prove the first assertion, assume, without loss in generality, that $f(\mathbf{x}_0) < f(\mathbf{x}_1)$, choose γ for $\mathbf{x}_0$ and $\mathbf{x}_1$, and set $u(t) = f\big(\gamma(t)\big)$. Then u is continuous on $[0,1]$, $u(0) = f(\mathbf{x}_0)$, and $u(1) = f(\mathbf{x}_1)$. Hence, by Theorem 1.3.6, for each $y \in \big[f(\mathbf{x}_0), f(\mathbf{x}_1)\big]$ there exists a $t \in [0,1]$ such that $u(t) = y$, and so we can take $\mathbf{x} = \gamma(t)$.

Given the first part, the second part follows immediately from the fact that f achieves both its minimum and maximum values on K. □

4.2 Differentiable Functions in Higher Dimensions

In view of the discussion of differentiable functions on $\mathbb{C}$, it should come as no surprise that we say that a $\mathbb{C}$-valued function f on an open set G in $\mathbb{R}^N$ is *differentiable* at $\mathbf{x} \in G$ in the direction $\boldsymbol{\xi} \in \mathbb{R}^N$ if the limit

$$\partial_{\boldsymbol{\xi}} f(\mathbf{x}) \equiv \lim_{t\to 0} \frac{f(\mathbf{x}+t\boldsymbol{\xi}) - f(\mathbf{x})}{t}$$

exists in $\mathbb{C}$, in which case $\partial_{\boldsymbol{\xi}} f(\mathbf{x})$ is called the *directional derivative* of f in the direction $\boldsymbol{\xi}$. When f is differentiable at $\mathbf{x}$ in all directions, we say that it is *differentiable* there. Now let $(\mathbf{e}_1, \dots, \mathbf{e}_N)$ be the *standard basis* in $\mathbb{R}^N$. That is, $\mathbf{e}_j = (\delta_{1,j}, \dots, \delta_{N,j})$, where the *Kronecker delta* $\delta_{i,j}$ equals 1 if $i = j$ and 0 if $i \neq j$. Clearly, $\boldsymbol{\xi} = \sum_{j=1}^N (\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N} \mathbf{e}_j$ and $|\boldsymbol{\xi}|^2 = \sum_{j=1}^N (\boldsymbol{\xi}, \mathbf{e}_j)^2_{\mathbb{R}^N}$ for every $\boldsymbol{\xi} \in \mathbb{R}^N$. Next suppose f is an $\mathbb{R}$-valued function that is differentiable at every point in G in each direction $\mathbf{e}_j$. If $\partial_{\mathbf{e}_j} f$ is continuous at $\mathbf{x}$ for each $1 \le j \le N$, then, for any $\boldsymbol{\xi} \in \mathbb{R}^N$,

$$\frac{f(\mathbf{x}+t\boldsymbol{\xi}) - f(\mathbf{x})}{t} = \sum_{j=1}^N \frac{f\big(\mathbf{x}_j(t) + t(\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N} \mathbf{e}_j\big) - f\big(\mathbf{x}_j(t)\big)}{t},$$

where

$$\mathbf{x}_j(t) = \begin{cases} \mathbf{x} + t\sum_{i=1}^{j-1} (\boldsymbol{\xi}, \mathbf{e}_i)_{\mathbb{R}^N} \mathbf{e}_i & \text{if } 2 \le j \le N \\ \mathbf{x} & \text{if } j = 1. \end{cases}$$

By Theorem 1.8.1, for each $1 \le j \le N$ there exists a $\tau_{j,t}$ between 0 and t such that

$$\frac{f\big(\mathbf{x}_j(t) + t(\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N} \mathbf{e}_j\big) - f\big(\mathbf{x}_j(t)\big)}{t} = (\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N} \partial_{\mathbf{e}_j} f\big(\mathbf{x}_j(t) + \tau_{j,t}(\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N} \mathbf{e}_j\big).$$

Hence, by Schwarz's inequality,

$$\left|\frac{f\big(\mathbf{x}+t\boldsymbol{\xi}\big)-f(\mathbf{x})}{t}-\sum_{j=1}^N(\boldsymbol{\xi},\mathbf{e}_j)_{\mathbb{R}^N}\partial_{\mathbf{e}_j}f(\mathbf{x})\right|$$
$$\le|\boldsymbol{\xi}|\left(\sum_{j=1}^N\left|\partial_{\mathbf{e}_j}f\big(\mathbf{x}_j(t)+\tau_{j,t}(\boldsymbol{\xi},\mathbf{e}_j)_{\mathbb{R}^N}\mathbf{e}_j\big)-\partial_{\mathbf{e}_j}f(\mathbf{x})\right|^2\right)^{\frac{1}{2}}$$

and so, for each $R\in(0,\infty)$,

$$\lim_{t\to 0}\sup_{|\boldsymbol{\xi}|\le R}\left|\frac{f\big(\mathbf{x}+t\boldsymbol{\xi}\big)-f(\mathbf{x})}{t}-\sum_{j=1}^N(\boldsymbol{\xi},\mathbf{e}_j)_{\mathbb{R}^N}\partial_{\mathbf{e}_j}f(\mathbf{x})\right|=0. \tag{4.2.1}$$

In particular, f is differentiable at $\mathbf{x}$ and

$$\partial_{\boldsymbol{\xi}}f(\mathbf{x})=\sum_{j=1}^N(\boldsymbol{\xi},\mathbf{e}_j)_{\mathbb{R}^N}\partial_{\mathbf{e}_j}f(\mathbf{x}), \tag{4.2.2}$$

which, as (4.2.1) makes clear, means that f is continuous at $\mathbf{x}$. (See Exercise 4.13 for an example that shows that (4.2.2) will not hold in general unless the $\partial_{\mathbf{e}_j}f$'s are continuous at $\mathbf{x}$.) When f is $\mathbb{C}$-valued, one can reach the same conclusions by applying the preceding to its real and imaginary parts.

When f is differentiable at each point in an open set G, we say that it is differentiable on G, and when it is differentiable on G and $\partial_{\boldsymbol{\xi}}f$ is continuous for all $\boldsymbol{\xi}\in\mathbb{R}^N$, then we say that it is continuously differentiable on G. In view of (4.2.2), it is obvious that f is continuously differentiable on G if $\partial_{\mathbf{e}_j}f$ is continuous for each $1\le j\le N$. When $f:G\longrightarrow\mathbb{R}$ is differentiable at $\mathbf{x}$, it is often convenient to introduce its *gradient*

$$\nabla f(\mathbf{x})\equiv\big(\partial_{\mathbf{e}_1}f(\mathbf{x}),\dots,\partial_{\mathbf{e}_N}f(\mathbf{x})\big)\in\mathbb{R}^N$$

there. For example, if f is continuously differentiable, then (4.2.2) can be written as $\partial_{\boldsymbol{\xi}}f(\mathbf{x})=\big(\nabla f(\mathbf{x}),\boldsymbol{\xi}\big)_{\mathbb{R}^N}$. Finally, given $f:G\longrightarrow\mathbb{C}$ and $n\ge 2$, we say that f is *n-times differentiable* on G if, for all $\boldsymbol{\xi}_1,\dots,\boldsymbol{\xi}_{n-1}\in\mathbb{R}^N$, f is differentiable on G and $\partial_{\boldsymbol{\xi}_m}\cdots\partial_{\boldsymbol{\xi}_1}f$ is differentiable on G for each $1\le m\le n-1$.

Just as in one dimension, derivatives simplify the search for places where an $\mathbb{R}$-valued function achieves its extreme values (i.e., either a maximum or minimum value). Namely, suppose that an $\mathbb{R}$-valued function f on a non-empty open set G achieves its maximum value at a point $\mathbf{x}\in G$. If f is differentiable at $\mathbf{x}$, then

$$\pm\partial_{\boldsymbol{\xi}} f(\mathbf{x}) = \lim_{t\searrow 0} \frac{f(\mathbf{x}\pm t\boldsymbol{\xi}) - f(\mathbf{x})}{t} \le 0 \quad \text{for all } \boldsymbol{\xi}\in\mathbb{R}^N,$$

and so $\partial_{\boldsymbol{\xi}} f(\mathbf{x}) = 0$. The same reasoning applies to minima and shows that $\partial_{\boldsymbol{\xi}} f(\mathbf{x}) = 0$ at an $\mathbf{x}$ where f achieves its minimum value. Hence, when looking for points at which f achieves extreme values, one need only look at points where its derivatives vanishes in all directions. In particular, when f is continuously differentiable, this means that, when searching for the places at which f achieves its extreme values, one can restrict one's attention to those $\mathbf{x}$ at which $\nabla f(\mathbf{x}) = \mathbf{0}$. This is the form that *the first derivative test* takes in higher dimensions. Next suppose that f is twice continuously differentiable on G and that it achieves its maximum value at a point $\mathbf{x}\in G$. Then (cf. Exercise 1.14)

$$\partial_{\boldsymbol{\xi}}^2 f(\mathbf{x}) = \lim_{t\to 0} \frac{f(\mathbf{x}+t\boldsymbol{\xi}) + f(\mathbf{x}-t\boldsymbol{\xi}) - 2f(\mathbf{x})}{t^2} \le 0 \quad \text{for all } \boldsymbol{\xi}\in\mathbb{R}^N.$$

Similarly, $\partial_{\boldsymbol{\xi}}^2 f(\mathbf{x}) \ge 0$ if f achieves its minimum value at $\mathbf{x}$, which is the *second derivative test* in higher dimensions.

Higher derivatives of functions on $\mathbb{R}^N$ bring up an interesting question. Namely, given $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ and a function f that is twice differentiable at $\mathbf{x}$, what is the relationship between $\partial_{\boldsymbol{\eta}}\partial_{\boldsymbol{\xi}} f(\mathbf{x})$ and $\partial_{\boldsymbol{\xi}}\partial_{\boldsymbol{\eta}} f(\mathbf{x})$?

Theorem 4.2.1 *Let $f : G \longrightarrow \mathbb{C}$ be a twice differentiable function on an open $G \ne \emptyset$. If $\boldsymbol{\xi}, \boldsymbol{\eta}\in\mathbb{R}^N$ and both $\partial_{\boldsymbol{\xi}}\partial_{\boldsymbol{\eta}} f$ and $\partial_{\boldsymbol{\eta}}\partial_{\boldsymbol{\xi}} f$ are continuous at the point $\mathbf{x}\in G$, then $\partial_{\boldsymbol{\xi}}\partial_{\boldsymbol{\eta}} f(\mathbf{x}) = \partial_{\boldsymbol{\eta}}\partial_{\boldsymbol{\xi}} f(\mathbf{x})$.*

Proof First observe that it suffices to treat the case when f is $\mathbb{R}$-valued. Second, by replacing f with

$$\mathbf{y} \rightsquigarrow f(\mathbf{x}+\mathbf{y}) - f(\mathbf{x}) - \sum_{j=1}^N (\mathbf{y},\mathbf{e}_j)\partial_{\mathbf{e}_j} f(\mathbf{x}),$$

we can reduce to the case when $\mathbf{x} = \mathbf{0}$ and $f(\mathbf{0}) = \partial_{\boldsymbol{\xi}} f(\mathbf{0}) = \partial_{\boldsymbol{\eta}} f(\mathbf{0}) = 0$. Thus, we will proceed under these assumptions.

Using Taylor's theorem, we know that, for small enough $t \ne 0$,

$$\begin{aligned} f(t\boldsymbol{\xi}+t\boldsymbol{\eta}) &= f(t\boldsymbol{\xi}) + t\partial_{\boldsymbol{\eta}} f(t\boldsymbol{\xi}) + \frac{t^2}{2}\partial_{\boldsymbol{\eta}}^2 f(t\boldsymbol{\xi}+\theta_t\boldsymbol{\eta}) \\ &= \frac{t^2}{2}\partial_{\boldsymbol{\xi}}^2 f(\theta_t'\boldsymbol{\xi}) + t^2\partial_{\boldsymbol{\xi}}\partial_{\boldsymbol{\eta}} f(\theta_t''\boldsymbol{\xi}) + \frac{t^2}{2}\partial_{\boldsymbol{\eta}}^2 f(t\boldsymbol{\xi}+\theta_t\boldsymbol{\eta}) \end{aligned}$$

for some choice of $\theta_t,\ \theta_t',\ \theta_t''$ lying between 0 and t. Similarly,

$$f(t\boldsymbol{\xi}+t\boldsymbol{\eta}) = \frac{t^2}{2}\partial_{\boldsymbol{\eta}}^2 f(\omega_t'\boldsymbol{\eta}) + t^2\partial_{\boldsymbol{\eta}}\partial_{\boldsymbol{\xi}} f(\omega_t''\boldsymbol{\eta}) + \frac{t^2}{2}\partial_{\boldsymbol{\xi}}^2 f(\omega_t\boldsymbol{\xi}+t\boldsymbol{\eta})$$

for some choice of $\omega_t,\ \omega_t',\ \omega_t''$ lying between 0 and t. Hence,

$$\begin{aligned}\tfrac{1}{2}\partial_{\boldsymbol{\xi}}^2 f(\theta_t''\boldsymbol{\xi}) + \partial_{\boldsymbol{\xi}}\partial_{\boldsymbol{\eta}} f(\theta_t''\boldsymbol{\xi}) + \tfrac{1}{2}\partial_{\boldsymbol{\eta}}^2 f(t\boldsymbol{\xi}+\theta_t\boldsymbol{\eta})\\ = \tfrac{1}{2}\partial_{\boldsymbol{\eta}}^2 f(\omega_t'\boldsymbol{\eta}) + \partial_{\boldsymbol{\eta}}\partial_{\boldsymbol{\xi}} f(\omega_t''\boldsymbol{\eta}) + \tfrac{1}{2}\partial_{\boldsymbol{\xi}}^2 f(\omega_t\boldsymbol{\xi}+t\boldsymbol{\eta}),\end{aligned}$$

and, after letting $t \to 0$, we arrive at $\partial_{\boldsymbol{\xi}}\partial_{\boldsymbol{\eta}} f(\mathbf{0}) = \partial_{\boldsymbol{\eta}}\partial_{\boldsymbol{\xi}} f(\mathbf{0})$. □

The result in Theorem 4.2.1 means that if f is n times continuously differentiable and $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n \in \mathbb{R}^N$, then $\partial_{\boldsymbol{\xi}_{\pi(1)}} \cdots \partial_{\boldsymbol{\xi}_{\pi(n)}} f$ is the same for all permutations π of $\{1, \dots, n\}$. This fact can be useful in many calculations, since a judicious choice of the order in which one performs the derivatives can greatly simplify the task. Another important application of this observation combined with (4.2.2) is the following. Given $\mathbf{m} = (m_1, \dots, m_N) \in \mathbb{N}^N$, define $\|\mathbf{m}\| = \sum_{j=1}^N m_j$, $\mathbf{m}! = \prod_{j=1}^N (m_j!)$, $\boldsymbol{\xi}^{\mathbf{m}} = \prod_{j=1}^N (\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N}^{m_j}$ for $\boldsymbol{\xi} \in \mathbb{R}^N$, and $\partial^{\mathbf{m}} f = \partial_{\mathbf{e}_1}^{m_1} \cdots \partial_{\mathbf{e}_N}^{m_N} f$ for $\|\mathbf{m}\|$ times differentiable f's, where $\partial_{\boldsymbol{\xi}}^0 f \equiv f$ for any $\boldsymbol{\xi}$. Then, for any $m \ge 0$, $\boldsymbol{\xi} \in \mathbb{R}^N$, and m times continuously differentiable f,

$$\partial_{\boldsymbol{\xi}}^m f = \sum_{\|\mathbf{m}\|=m} \frac{m!}{\mathbf{m}!} \boldsymbol{\xi}^{\mathbf{m}} \partial^{\mathbf{m}} f. \tag{4.2.3}$$

To check (4.2.3), use (4.2.2) to see that

$$\partial_{\boldsymbol{\xi}}^m f = \sum_{(j_1,\dots,j_m)\in\{1,\dots,N\}^m} \left(\prod_{i=1}^m (\boldsymbol{\xi}, \mathbf{e}_{j_i})_{\mathbb{R}^N}\right) \partial_{\mathbf{e}_{j_1}} \cdots \partial_{\mathbf{e}_{j_m}} f.$$

Next, given $\mathbf{m}$ with $\|\mathbf{m}\| = m$, let $S(\mathbf{m})$ be the set of $(j_1, \dots, j_m) \in \{1, \dots, N\}^m$ such that, for each $1 \le j \le N$, exactly m_j of the j_i's are equal to j, and observe that, for all $(j_1, \dots, j_m) \in S(\mathbf{m})$, $\prod_{i=1}^m (\boldsymbol{\xi}, \mathbf{e}_{j_i})_{\mathbb{R}^N} = \boldsymbol{\xi}^{\mathbf{m}}$ and $\partial_{\mathbf{e}_{j_1}} \cdots \partial_{\mathbf{e}_{j_m}} = \partial^{\mathbf{m}} f$. Thus

$$\partial_{\boldsymbol{\xi}}^m f = \sum_{\|\mathbf{m}\|=m} \operatorname{card}\big(S(\mathbf{m})\big) \boldsymbol{\xi}^{\mathbf{m}} \partial^{\mathbf{m}} f.$$

Finally, to compute the number of elements in $S(\mathbf{m})$, let $(j_1^0, \dots, j_m^0)$ be the element of $S(\mathbf{m})$ in which the first m_1 entries are 1, the next m_2 are 2, etc. Then all the other elements of $S(\mathbf{m})$ can be obtained by at least one of the $m!$ permutations of the indices of $(j_1^0, \dots, j_m^0)$. However, $\mathbf{m}!$ of these permutations will result in the same element, and so $S(\mathbf{m})$ has $\frac{m!}{\mathbf{m}!}$ elements.

If f is a twice continuously differentiable function, define the *Hessian* $Hf(\mathbf{x})$ of f at $\mathbf{x}$ to be the mapping of $\mathbb{R}^N$ into $\mathbb{R}^N$ given by

$$Hf(\mathbf{x})\boldsymbol{\xi} = \sum_{i=1}^N \left(\sum_{j=1}^N \partial_{\mathbf{e}_j}\partial_{\mathbf{e}_j} f(\mathbf{x})(\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N}\right) \mathbf{e}_i.$$

Then $\partial_{\boldsymbol{\xi}}\partial_{\boldsymbol{\eta}} f(\mathbf{x}) = \big(\boldsymbol{\xi}, Hf(\mathbf{x})\boldsymbol{\eta}\big)_{\mathbb{R}^N}$, and, from Theorem 4.2.1, we see that $Hf(\mathbf{x})$ is symmetric in the sense that $\big(\boldsymbol{\xi}, Hf(\mathbf{x})\boldsymbol{\eta}\big)_{\mathbb{R}^N} = \big(\boldsymbol{\eta}, Hf(\mathbf{x})\boldsymbol{\xi}\big)_{\mathbb{R}^N}$. In this notation, the second derivative test becomes the statement that f achieves a maximum at $\mathbf{x}$ only if $\big(\boldsymbol{\xi}, Hf(\mathbf{x})\boldsymbol{\xi}\big)_{\mathbb{R}^N} \le 0$ for all $\boldsymbol{\xi} \in \mathbb{R}^N$. There are many results (cf. Exercise 4.14) in linear algebra that provide criteria with which to determine when this kind of inequality holds.

Theorem 4.2.2 (Taylor's Theorem for $\mathbb{R}^N$) *Let f be an $(n+1)$ times differentiable $\mathbb{R}$-valued function on the open ball $B(\mathbf{x}, r)$ in $\mathbb{R}^N$. Then for each $\mathbf{y} \in B(\mathbf{x}, r)$ there exists a $\theta \in (0, 1)$ such that*

$$f(\mathbf{y}) = \sum_{\|\mathbf{m}\| \le n} \frac{\partial^{\mathbf{m}} f(\mathbf{x})}{\mathbf{m}!}(\mathbf{y}-\mathbf{x})^{\mathbf{m}} + \frac{\partial^{n+1}_{\mathbf{y}-\mathbf{x}} f\big((1-\theta)\mathbf{x}+\theta\mathbf{y}\big)}{(n+1)!}.$$

Moreover, if f is $(n+1)$ times continuously differentiable, then the last term can be replaced by

$$\sum_{\|\mathbf{m}\| = n+1} \frac{\partial^{\mathbf{m}} f\big((1-\theta)\mathbf{x}+\theta\mathbf{y}\big)}{\mathbf{m}!}(\mathbf{y}-\mathbf{x})^{\mathbf{m}}.$$

Proof Set $\boldsymbol{\xi} = \mathbf{y} - \mathbf{x}$. Then, by Theorem 1.8.2 applied to $t \rightsquigarrow f(\mathbf{x} + t\boldsymbol{\xi})$,

$$f(\mathbf{y}) = f(\mathbf{x}+\boldsymbol{\xi}) = \sum_{m=0}^{n} \frac{\partial^m_{\boldsymbol{\xi}} f(\mathbf{x})}{m!} + \frac{\partial^{n+1}_{\boldsymbol{\xi}} f(\mathbf{x}+\theta\boldsymbol{\xi})}{(n+1)!}$$

for some $\theta \in (0, 1)$. Since f being $(n+1)$ times differentiable on $B(\mathbf{x}, r)$ implies that it is n times continuously differentiable there, (4.2.3) implies the first assertion. As for the second assertion, when f is $(n+1)$ times continuously differentiable, another application of (4.2.3) yields the desired result. □

4.3 Arclength of and Integration Along Paths

One of the many applications of integration is to the computation of lengths of paths in $\mathbb{R}^N$. That is, given a continuous path $\mathbf{p} : [a, b] \longrightarrow \mathbb{R}^N$, we want a way to compute its length. If $\mathbf{p}$ is linear, in the sense that $\mathbf{p}(t) - \mathbf{p}(a) = t\mathbf{v}$ for some $\mathbf{v} \in \mathbb{R}^N$, then it is the trajectory of a particle that is moving in a straight line with constant velocity $\mathbf{v}$. Hence its length should be the distance that the particle traveled, namely, $(b-a)|\mathbf{v}|$. Further, when $\mathbf{p}$ is piecewise linear, in the sense that it can be obtained by concatenating a finite number of linear paths, then its length should be the sum of the lengths of the linear paths out of which it is made. With this in mind, we adopt the following procedure for computing the length of a general path. Given a finite cover $\mathcal{C}$ of $[a, b]$ by non-overlapping closed intervals, define $\mathbf{p}_{\mathcal{C}}$ to be the piecewise linear path which, for each $I \in \mathcal{C}$, is linear on I and agrees with $\mathbf{p}$ at the endpoints. That is, if $I = [a_I, b_I]$ where $a_I < b_I$, then

$$\mathbf{p}_{\mathcal{C}}(t) = \frac{b_I - t}{b_I - a_I}\mathbf{p}(a_I) + \frac{t - a_I}{b_I - a_I}\mathbf{p}(b_I) \quad \text{for } t \in I.$$

Clearly $\mathbf{p}_{\mathcal{C}}$ tends uniformly fast to $\mathbf{p}$ as $\|\mathcal{C}\| \to 0$, and its length is

$$\sum_{I \in \mathcal{C}} |\Delta_I \mathbf{p}| \quad \text{where } \Delta_I \mathbf{p} = \mathbf{p}(b_I) - \mathbf{p}(a_I).$$

Thus a good definition of the length of $\mathbf{p}$ would be

$$L_{\mathbf{p}}\big([a, b]\big) = \lim_{\|\mathcal{C}\| \to 0} \sum_{I \in \mathcal{C}} |\Delta_I \mathbf{p}|. \tag{4.3.1}$$

However, we first have to check that this limit exists.

Lemma 4.3.1 *Referring to the preceding, the limit in (4.3.1) exists and is equal to*

$$\sup_{\mathcal{C}} \sum_{I \in \mathcal{C}} |\Delta_I \mathbf{p}|.$$

Proof To prove this result, what we have to show is that for each $\mathcal{C}$ and $\epsilon > 0$ there is a $\delta > 0$ such that

$$(*) \qquad \sum_{I \in \mathcal{C}} |\Delta_I \mathbf{p}| \le \sum_{I' \in \mathcal{C}'} |\Delta_{I'} \mathbf{p}| + \epsilon \quad \text{if } \|\mathcal{C}'\| < \delta.$$

Without loss in generality, we will assume that $r = \min\{|I| : I \in \mathcal{C}\} > 0$. Take $0 < \delta < r$ so that $|t - s| < \delta \implies |\mathbf{p}(t) - \mathbf{p}(s)| < \frac{\epsilon}{2n}$, where n is the number of elements in $\mathcal{C}$. If $\mathcal{C}'$ with $\|\mathcal{C}'\| < \delta$ is given, then, by the triangle inequality,

$$\sum_{I \in \mathcal{C}} |\Delta_I \mathbf{p}| \le \sum_{I \in \mathcal{C}} \sum_{I' \in \mathcal{C}'} |\Delta_{I \cap I'} \mathbf{p}|.$$

Clearly,

$$\sum_{I \in \mathcal{C}} \sum_{\substack{I' \in \mathcal{C}' \\ I' \subseteq I}} |\Delta_{I \cap I'} \mathbf{p}| = \sum_{I' \in \mathcal{C}'} \sum_{\substack{I \in \mathcal{C} \\ I \supseteq I'}} |\Delta_{I \cap I'} \mathbf{p}| \le \sum_{I' \in \mathcal{C}'} |\Delta_{I'} \mathbf{p}|.$$

At the same time, since, for any $I \in \mathcal{C}$, there are most two $I' \in \mathcal{C}'$ for which $\Delta_{I' \cap I}\mathbf{p} \neq 0$ but $I' \nsubseteq I$,

$$\sum_{I \in \mathcal{C}} \sum_{\substack{I' \in \mathcal{C}' \\ I' \cap I \neq \emptyset \\ I' \nsubseteq I}} |\Delta_{I \cap I'}\mathbf{p}| \leq 2n \max\{|\mathbf{p}(t) - \mathbf{p}(s)| : |t-s| < \delta\} < \epsilon,$$

and so $(*)$ holds. □

Now that we know that the limit on the right hand side of (4.3.1) exists, we will say that the *arclength* of $\mathbf{p}$ is the number $L_{\mathbf{p}}\big([a,b]\big)$. Notice that $L_{\mathbf{p}}\big([a,b]\big)$ need not be finite. For example, consider the function $f : [0,1] \longrightarrow [-1,1]$ defined by

$$f(t) = \begin{cases} (n+1)^{-1} \sin(4^{n+1}\pi t) & \text{if } 1-2^{-n} \leq t < 1-2^{-n-1} \text{ for } n \geq 0 \\ 0 & \text{if } t = 1, \end{cases}$$

and set $\mathbf{p}(t) = \big(t, f(t)\big)$. Then $\mathbf{p}$ is a continuous path in $\mathbb{R}^2$. However, because, for each $n \geq 0$, f takes the values $-(n+1)^{-1}$ and $(n+1)^{-1}$ 2^n times in the interval $[1-2^{-n}, 1-2^{-n-1}]$, it is easy to check that $L_{\mathbf{p}}\big([0,1]\big) = \infty$.

We show next that length is additive in the sense that

$$L_{\mathbf{p}}\big([a,b]\big) = L_{\mathbf{p}}\big([a,c]\big) + L_{\mathbf{p}}\big([c,b]\big) \quad \text{for } c \in (a,b). \tag{4.3.2}$$

To this end, first observe that if $\mathcal{C}_1$ is a cover for $[a,c]$ and $\mathcal{C}_2$ is a cover for $[c,b]$, then $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$ is a cover for $[a,b]$ and therefore

$$L_{\mathbf{p}}\big([a,b]\big) \geq \sum_{I \in \mathcal{C}_1} |\Delta_I \mathbf{p}| + \sum_{I \in \mathcal{C}_2} |\Delta_I \mathbf{p}|.$$

Hence, the left hand side of (4.3.2) dominates the right hand side. To prove the opposite inequality, suppose that $\mathcal{C}$ is a cover for $[a,b]$. If $c \notin \operatorname{int}(I)$ for any $I \in \mathcal{C}$, then $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$, where $\mathcal{C}_1 = \{I \in \mathcal{C} : I \subseteq [a,c]\}$ and $\mathcal{C}_2 = \{I \in \mathcal{C} : I \subseteq [c,b]\}$. In addition, $\mathcal{C}_1$ is a cover for $[a,c]$ and $\mathcal{C}_2$ is a cover for $[c,b]$. Hence

$$\sum_{I \in \mathcal{C}} |\Delta_I \mathbf{p}| = \sum_{I \in \mathcal{C}_1} |\Delta_I \mathbf{p}| + \sum_{I \in \mathcal{C}_2} |\Delta_I \mathbf{p}| \leq L_{\mathbf{p}}\big([a,c]\big) + L_{\mathbf{p}}\big([c,b]\big)$$

in this case. Next assume that $c \in \operatorname{int}(I)$ for some $I \in \mathcal{C}$. Because the elements of $\mathcal{C}$ are non-overlapping, there is precisely one such I, and we will use J to denote it. If $J_- = J \cap [a,c]$ and $J_+ = J \cap [c,b]$, then

$$\mathcal{C}_1 = \{I \in \mathcal{C} : I \subseteq [a,c]\} \cup \{J_-\} \text{ and } \mathcal{C}_2 = \{I \in \mathcal{C} : I \subseteq [c,b]\} \cup \{J_+\}$$

are non-overlapping covers of $[a, c]$ and $[c, b]$. Since $|\Delta_J \mathbf{p}| \le |\Delta_{J_-}\mathbf{p}| + |\Delta_{J_+}\mathbf{p}|$, it follows that

$$L_{\mathbf{p}}\big([a,c]\big) + L_{\mathbf{p}}\big([c,b]\big) \ge \sum_{I\in\mathcal{C}_1} |\Delta_I \mathbf{p}| + \sum_{I\in\mathcal{C}_2} |\Delta_I \mathbf{p}| \ge \sum_{I\in\mathcal{C}} |\Delta_I \mathbf{p}|,$$

and this completes the proof of (4.3.2).

We now have enough information to describe what we will mean by integration of functions along a path of finite length. Suppose that $\mathbf{p} : [a, b] \longmapsto \mathbb{R}^N$ is a continuous path for which $L_{\mathbf{p}}\big([a, b]\big) < \infty$, and let $f : \mathbf{p}([a, b]) \longrightarrow \mathbb{C}$ be a bounded function. Given a finite cover $\mathcal{C}$ of $[a, b]$ by non-overlapping, closed intervals and an associated choice function Ξ, consider the quantity

$$\sum_{I\in\mathcal{C}} (f\circ\mathbf{p})\big(\Xi(I)\big)L_{\mathbf{p}}(I).$$

If, as $\|\mathcal{C}\| \to 0$, these quantities converge in $\mathbb{R}$ to a limit, then that limit is what we will call the *integral* of f along $\mathbf{p}$. Because of (4.3.2), the problem of determining when such a limit exists and of understanding its properties when it does can be solved using Riemann–Stieltjes integration. Indeed, define $F_{\mathbf{p}} : [a, b] \longrightarrow [0, \infty)$ by

$$F_{\mathbf{p}}(t) = L_{\mathbf{p}}\big([a, t]\big) \quad \text{for } t \in [a, b], \tag{4.3.3}$$

Clearly $F_{\mathbf{p}}$ is a non-decreasing function,[1] and, by (4.3.2), $L_{\mathbf{p}}(I) = \Delta_I F_{\mathbf{p}}$. Thus a limit will exist precisely when $f \circ \mathbf{p}$ is Riemann–Stieltjes integrable with respect to $F_{\mathbf{p}}$ on $[a, b]$, in which case

$$\lim_{\|\mathcal{C}\|\to 0} \sum_{I\in\mathcal{C}} (f\circ\mathbf{p})\big(\Xi(I)\big)L_{\mathbf{p}}(I) = \int_a^b (f\circ\mathbf{p})(t)\, dF_{\mathbf{p}}(t).$$

The following lemma provides an important computational tool.

Lemma 4.3.2 *If* $\mathbf{p} : [a, b] \longrightarrow \mathbb{R}^N$ *is continuous path of finite length which is continuously differentiable on* (a, b) *and*[2]

$$\dot{\mathbf{p}}(t) = \big(p_1'(t), \ldots, p_N'(t)\big) \quad \textit{for } t \in (a, b),$$

[1]Although we will not prove it, $F_{\mathbf{p}}$ is also continuous. For those who know about such things, one way to prove this is to observe that, because $\mathbf{p}$ has finite length, each coordinate p_j is a continuous function of bounded variation. Hence the variation of p_j over an interval is a continuous function of the endpoints of the interval, and from this it is easy to check the continuity of $F_{\mathbf{p}}$. See Exercise 1.2.22 in my book *Essentials of Integration Theory of Analysts*, Springer-Verlag GTM 262, for more details.

[2]The use of a "dot" to denote the time derivative of a vector valued function goes back to Newton.

then

$$\int_a^b (f \circ \mathbf{p})(t)\, dF_{\mathbf{p}}(t) = \int_{(a,b)} (f \circ \mathbf{p})(t)|\dot{\mathbf{p}}(t)|\, dt$$

for continuous functions $f : [a, b] \longrightarrow \mathbb{R}$.

Proof By (3.5.3), all that we have to show is that $F_{\mathbf{p}}$ is differentiable on (a, b) and that $|\dot{\mathbf{p}}|$ is its derivative there. Hence, by the Fundamental of Calculus, it suffices to show that

$$(*) \qquad F_{\mathbf{p}}(t) - F_{\mathbf{p}}(s) = \int_s^t |\dot{\mathbf{p}}(\tau)|\, d\tau \quad \text{for } a < s < t < b.$$

Given $a < s < t < b$, let $\mathcal{C}$ be a non-overlapping cover for $[s, t]$. Then, by Theorem 1.8.1, for each $I = [a_I, b_I] \in \mathcal{C}$ there exist $\xi_{1,I}, \dots, \xi_{N,I} \in I$ such that

$$\begin{aligned} \Delta_I \mathbf{p} &= \big(p_1'(\xi_{1,I}), \dots, p_N'(\xi_{N,I})\big)|I| \\ &= \big(p_1'(a_I), \dots, p_N'(a_I)\big)|I| + \big(p_1'(\xi_{1,I}) - p_1'(a_I), \dots, p_N'(\xi_{N,I}) - p_N'(a_I)\big)|I|. \end{aligned}$$

Hence, if

$$\omega(\delta) = \max_{1 \le j \le N} \sup\big\{|p_j'(\tau) - p_j'(\sigma)| : s \le \sigma < \tau \le t \;\&\; \tau - \sigma \le \delta\big\},$$

then, by the triangle and Schwarz's inequalities,

$$\Big||\Delta_I \mathbf{p}| - |\dot{\mathbf{p}}(a_I)||I|\Big| \le \big|\Delta_I \mathbf{p} - |I|\dot{\mathbf{p}}(a_I)\big| \le N^{\frac{1}{2}}\omega(\|\mathcal{C}\|)|I|,$$

and so

$$F_{\mathbf{p}}(t) - F_{\mathbf{p}}(s) = \lim_{\|\mathcal{C}\| \to 0} \sum_{I \in \mathcal{C}} |\Delta_I \mathbf{p}| = \lim_{\|\mathcal{C}\| \to 0} \sum_{I \in \mathcal{C}} |\dot{\mathbf{p}}(a_I)||I| = \int_s^t |\dot{\mathbf{p}}(\tau)|\, d\tau.$$

□

4.4 Integration on Curves

The reader might well be wondering why we defined integrals along paths the way that we did instead of simply as $\int_a^b (f \circ \mathbf{p})(t)\, dt$, but the explanation will become clear in this section (cf. Lemma 4.4.1).

We will say that a compact subset C of $\mathbb{R}^N$ is a continuously *parameterized curve* if $C = \mathbf{p}([a, b])$ for some one-to-one, continuous map $\mathbf{p} : [a, b] \longrightarrow \mathbb{R}^N$, called a *parameterization* of C. Given a continuous function $f : C \longrightarrow \mathbb{C}$, we want to know what meaning to give the *integral* $\int_C f(\mathbf{y})\, d\sigma_{\mathbf{y}}$ of f over C. Here the notation $d\sigma_{\mathbf{y}}$ is used to emphasize that the integral is being taken over C and is not a usual integral in Euclidean space, but one that is intrinsically defined on C.

One possibility that suggests itself is to take (cf. (4.3.3))

$$\int_C f(\mathbf{y})\, d\sigma_{\mathbf{y}} = \int_a^b f\big(\mathbf{p}(t)\big)\, dF_{\mathbf{p}}(t) \tag{4.4.1}$$

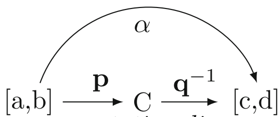
at least when $L_{\mathbf{p}}\big([a,b]\big) < \infty$. However, to justify defining integrals over C by (4.4.1), we have to make sure that the right hand side depends only on C in the sense that it is the same for all parameterizations of C.

Lemma 4.4.1 *Suppose that* $\mathbf{p} : [a,b] \longrightarrow C$ *and* $\mathbf{q} : [c,d] \longrightarrow C$ *are two parameterizations of* C. *Then* $L_{\mathbf{p}}\big([a,b]\big) = L_{\mathbf{q}}\big([c,d]\big)$, *and, if* $L_{\mathbf{p}}\big([a,b]\big) < \infty$, *then*

$$\int_a^b f\big(\mathbf{p}(t)\big)\, dF_{\mathbf{p}}(t) = \int_c^d f\big(\mathbf{q}(t)\big)\, dF_{\mathbf{q}}(t)$$

for all continuous functions $f : C \longrightarrow \mathbb{C}$.

Proof First observe that, because $\mathbf{q}$ is one-to-one onto C, it admits an inverse $\mathbf{q}^{-1}$ taking C onto $[c,d]$. Furthermore, $\mathbf{q}^{-1}$ is continuous, since if $\mathbf{x}_n \longrightarrow \mathbf{x}$ in C and $t_n = \mathbf{q}^{-1}(\mathbf{x}_n)$, then every limit point of the sequence $\{t_n : n \geq 1\}$ is mapped by $\mathbf{q}$ to $\mathbf{x}$ and therefore (cf. Exercise 4.2) $t_n \longrightarrow \mathbf{q}^{-1}(\mathbf{x})$. Now define $\alpha(s) = \mathbf{q}^{-1} \circ \mathbf{p}(s)$ for $s \in [a,b]$.



commutative diagram

Then, since $\mathbf{q}^{-1}$ and $\mathbf{p}$ are continuous, one-to-one, and onto, α is a continuous, one-to-one map of $[a,b]$ onto $[c,d]$. Thus, by Corollary 1.3.7, either $\alpha(a) = c$ and α is strictly increasing, or $\alpha(a) = d$ and α is strictly decreasing.

Assuming that $\alpha(a) = c$, we now show that

$$(*) \qquad L_{\mathbf{p}}(J) = L_{\mathbf{q}}\big(\alpha(J)\big) \quad \text{for each closed interval } J \subseteq [a,b].$$

To this end, let $\mathcal{C}$ be a cover of J by non-overlapping closed intervals I. Then $\mathcal{C}' = \big\{\alpha(I) : I \in \mathcal{C}\big\}$ is a non-overlapping cover of $\alpha(J)$ (cf. Theorem 1.3.6 and Exercise 4.6) by non-overlapping closed intervals, and $\|\mathcal{C}'\| \to 0$ as $\|\mathcal{C}\| \to 0$. Hence, since $\Delta_{\alpha(I)}\mathbf{q} = \Delta_I \mathbf{p}$,

$$L_{\mathbf{q}}\big(\alpha(J)\big) = \lim_{\|\mathcal{C}\| \to 0} \sum_{I \in \mathcal{C}} |\Delta_{\alpha(I)}\mathbf{q}| = \lim_{\|\mathcal{C}\| \to 0} \sum_{I \in \mathcal{C}} |\Delta_I \mathbf{p}| = L_{\mathbf{p}}(J).$$

Obviously, the first assertion follows immediately by taking $J = [a, b]$ in (∗). Further, if $L_{\mathbf{p}}\big([a, b]\big) < \infty$, (∗) implies that $F_{\mathbf{p}}(s) = F_{\mathbf{q}}\big(\alpha(s)\big)$. Now suppose that f is a continuous function on C. Given a cover $\mathcal{C}$ of $[a, b]$ and an associated choice function Ξ, set $\mathcal{C}' = \{\alpha(I) : I \in \mathcal{C}\}$ and $\Xi'\big(\alpha(I)\big) = \Xi(I)$ for $I \in \mathcal{C}$. Then

$$\mathfrak{R}(f|F_{\mathbf{q}}; \mathcal{C}', \Xi') = \sum_{I \in \mathcal{C}} f \circ \mathbf{q}\big(\Xi'(\alpha(I))\big)\Delta_{\alpha(I)} F_{\mathbf{q}} = \mathfrak{R}(f|F_{\mathbf{p}}; \mathcal{C}, \Xi),$$

and so the Riemann–Stieltjes integrals of f with respect to $F_{\mathbf{p}}$ and $F_{\mathbf{q}}$ are equal.

When $\alpha(a) = d$, set $\tilde{\mathbf{q}}(t) = \mathbf{q}(c + d - t)$. Since $\tilde{\mathbf{q}}^{-1} \circ \mathbf{p} = \tilde{\alpha}(s)$, where $\tilde{\alpha}(s) = \alpha(a + b - s)$, the preceding says that $L_{\mathbf{p}}\big([a, b]\big) = L_{\tilde{\mathbf{q}}}\big([c, d]\big)$ and, when $L_{\mathbf{p}}\big([a, b]\big) < \infty$,

$$\int_a^b f\big(\mathbf{p}(s)\big)\, dF_{\mathbf{p}}(s) = \int_c^d f\big(\tilde{\mathbf{q}}(t)\big)\, dF_{\tilde{\mathbf{q}}}(t).$$

Clearly $L_{\mathbf{q}}\big([c, d]\big) = L_{\tilde{\mathbf{q}}}\big([c, d]\big)$. Moreover, if $L_{\mathbf{p}}\big([a, b]\big) < \infty$, then $F_{\tilde{\mathbf{q}}}(t) = F_{\mathbf{q}}(b) - F_{\mathbf{q}}(a + b - t)$, and so, by (3.5.1),

$$\int_c^d f\big(\mathbf{q}(t)\big)\, dF_{\mathbf{q}}(t) = \int_c^d f\big(\tilde{\mathbf{q}}(t)\big)\, dF_{\tilde{\mathbf{q}}}(t).$$ □

In view of the preceding, it makes sense to say that the *length* of C is the length of a, and therefore any, parameterization of C. In addition, if C has finite length and $\mathbf{p}$ is a parameterization, then we can use (4.4.1) to define $\int_C f(\mathbf{y})\, d\sigma_{\mathbf{y}}$, and it makes no difference which parameterization we use. In many applications, the freedom to choose the parameterization when performing integrals along parameterized curves of finite length can greatly simplify computations. In particular, if C has finite length and admits a parameterization $\mathbf{p} : [a, b] \longrightarrow C$ that is continuously differentiable on (a, b), then, by Lemma 4.3.2,

$$\int_C f(\mathbf{y})\, d\sigma_{\mathbf{y}} = \int_a^b f\big(\mathbf{p}(t)\big)|\dot{\mathbf{p}}(t)|\, dt. \tag{4.4.2}$$

Here is an example that illustrates the point being made. Let C be the semi-circle $\{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x}| = 1 \;\&\; x_2 \geq 0\}$. Then there are two parameterizations that suggest themselves. The first is $t \in [-1, 1] \longmapsto \mathbf{p}(t) = \big(t, \sqrt{1 - t^2}\big) \in C$, and the second is $t \in [0, \pi] \longmapsto \mathbf{q}(t) = \big(\cos t, \sin t\big)$. Note that

$$F_{\mathbf{p}}(t) = \int_{(-1,t]} \sqrt{1 + \tfrac{\tau^2}{1-\tau^2}}\, d\tau = \int_{(-1,t]} \frac{1}{\sqrt{1 - \tau^2}}\, d\tau = \pi - \arccos t,$$

where $\arccos t$ is the $\theta \in [0, \pi]$ for which $\cos\theta = t$, and $F_{\mathbf{q}}(t) = t$ for $t \in [0, \pi]$. Hence, in this case, the equality between integrals along $\mathbf{p}$ and integrals along $\mathbf{q}$ comes from the change of variables $t = \cos s$ in the integral along $\mathbf{p}$.

Before closing this section, we need to discuss an important extension of these ideas. Namely, one often wants to deal with curves C that are best presented as the union of parameterized curves that are *non-overlapping* in the sense that if $\mathbf{p}$ on $[a, b]$ is a parameterization of one and $\mathbf{q}$ on $[c, d]$ is a parameterization of another, then $\mathbf{p}(s) \neq \mathbf{q}(t)$ for any $s \in (a, b)$ and $t \in (c, d)$. For example, the unit circle $\mathbb{S}^1(\mathbf{0}, 1) = \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x}| = 1\}$ cannot be parameterized because no continuous path $\mathbf{p}$ from a closed interval onto $\mathbb{S}^1(\mathbf{0}, 1)$ could be one-to-one. On the other hand, $\mathbb{S}^1(\mathbf{0}, 1)$ is the union of two non-overlapping parameterized curves: $\{\mathbf{x} \in \mathbb{S}^1(\mathbf{0}, 1) : x_2 \geq 0\}$ and $\{\mathbf{x} \in \mathbb{S}^1(\mathbf{0}, 1) : x_2 \leq 0\}$. The obvious way to integrate over such a curve is to sum the integrals over its parameterized parts. That is, if $C = C_1 \cup \cdots \cup C_M$, the C_m's being non-overlapping, parameterized curves, then one says that C is *piecewise parameterized*, takes its *length* to be the sum of the lengths of its parameterized components, and, when its length is finite, the *integral* of a continuous function f over C to be

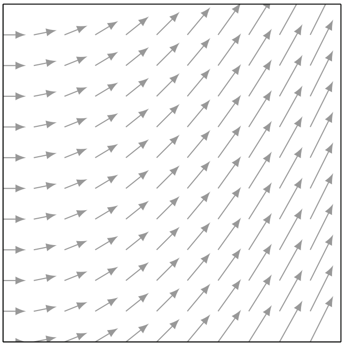
$$\int_C f(\mathbf{y})\, d\sigma_{\mathbf{y}} = \sum_{m=1}^{M} \int_{C_m} f(\mathbf{y})\, d\sigma_{\mathbf{y}},$$

the sum of the integrals of f over its components. Of course, one has to make sure that these definitions lead to the same answer for all decompositions of C into non-overlapping parameterized components. However, if $C'_1, \ldots, C'_{M'}$ is a second decomposition, then one can consider the decomposition whose components are of the form $C_m \cap C'_{m'}$ for m and m' corresponding to overlapping components. Using Lemma 4.4.1, one sees that both of the original decompositions give the same answers on the components of this third one, and from this it is an easy step to show that our definitions of length and integrals on C do not depend on the way in which C is decomposed into non-overlapping, parameterized curves.

4.5 Ordinary Differential Equations[3]

One way to think about a function $\mathbf{F} : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is as a *vector field*. That is, to each $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{F}$ assigns the vector $\mathbf{F}(\mathbf{x})$ at $\mathbf{x}$:

[3]The adjective "ordinary" is used to distinguish differential equations in which, as opposed to "partial" differential equations, the derivatives are all taken with respect to one variable.



vector field

and one can ask whether there exists a path $\mathbf{X} : \mathbb{R} \longrightarrow \mathbb{R}^N$ for which $\mathbf{F}\big(\mathbf{X}(t)\big)$ is the velocity of $\mathbf{X}(t)$ at each time $t \in \mathbb{R}$. That is, $\dot{\mathbf{X}}(t) \equiv \frac{d}{dt}\mathbf{X}(t) = \mathbf{F}\big(\mathbf{X}(t)\big)$. In fact, it is reasonable to hope that, under appropriate conditions on $\mathbf{F}$, for each $\mathbf{x} \in \mathbb{R}^N$ there will be precisely one path $t \rightsquigarrow \mathbf{X}(t, \mathbf{x})$ that passes through $\mathbf{x}$ at time $t = 0$ and has velocity $\mathbf{F}\big(\mathbf{X}(t, \mathbf{x})\big)$ at all times $t \in \mathbb{R}$. The most intuitive way to go about constructing such a path is to pretend that $\mathbf{F}\big(\mathbf{X}(t, \mathbf{x})\big)$ is constant during time intervals of length $\frac{1}{n}$. In other words, consider the path $t \rightsquigarrow \mathbf{X}_n(t, \mathbf{x})$ such that $\mathbf{X}_n(0, \mathbf{x}) = \mathbf{x}$ and

$$\dot{\mathbf{X}}_n(t, \mathbf{x}) = \mathbf{F}\big(\mathbf{X}_n\big(\tfrac{k}{n}, \mathbf{x}\big)\big) \quad \text{for} \quad \begin{cases} t \in \big[\frac{k}{n}, \frac{k+1}{n}\big) & \text{if } t \geq 0 \\ t \in \big(\frac{k-1}{n}, \frac{k}{n}\big] & \text{if } t < 0. \end{cases}$$

Equivalently, if

$$\lfloor t \rfloor_n \equiv \frac{\max\big\{k \in \mathbb{Z} : \frac{k}{n} \leq t\big\}}{n} \text{ and } \lceil t \rceil_n \equiv \frac{\min\big\{k \in \mathbb{Z} : \frac{k}{n} \geq t\big\}}{n},$$

then

$$\mathbf{X}_n(t, \mathbf{x}) = \begin{cases} \mathbf{X}_n\big(\lfloor t \rfloor_n, \mathbf{x}\big) + \mathbf{F}\big(\mathbf{X}_n\big(\lfloor t \rfloor_n, \mathbf{x}\big)\big(t - \lfloor t \rfloor_n\big) & \text{if } t > 0 \\ \mathbf{X}_n\big(\lceil t \rceil_n, \mathbf{x}\big) + \mathbf{F}\big(\mathbf{X}_n\big(\lceil t \rceil_n, \mathbf{x}\big)\big(t - \lceil t \rceil_n\big) & \text{if } t < 0, \end{cases}$$

which can be also written as

$$\mathbf{X}_n(t, \mathbf{x}) = \mathbf{x} + \begin{cases} \int_0^t \mathbf{F}\big(\mathbf{X}_n\big(\lfloor \tau \rfloor_n, \mathbf{x}\big)\big)\, d\tau & \text{if } t \geq 0 \\ -\int_0^{-t} \mathbf{F}\big(\mathbf{X}_n\big(-\lfloor -\tau \rfloor_n, \mathbf{x}\big)\, d\tau & \text{if } t < 0. \end{cases} \tag{4.5.1}$$

The hope is that, as $n \to \infty$, the paths $t \rightsquigarrow \mathbf{X}_n(t, \mathbf{x})$ will converge to a solution to

$$\mathbf{X}(t, \mathbf{x}) = \mathbf{x} + \int_0^t \mathbf{F}\big(\mathbf{X}(\tau, \mathbf{x})\big)\, d\tau, \tag{4.5.2}$$

which, by the Fundamental Theorem of Calculus, is that same as saying it is a solution to

$$\dot{\mathbf{X}}(t, \mathbf{x}) = \mathbf{F}\big(\mathbf{X}(t, \mathbf{x})\big) \quad \text{with } \mathbf{X}(0, \mathbf{x}) = \mathbf{x}. \tag{4.5.3}$$

Notice that $t \rightsquigarrow \mathbf{X}(-t, \mathbf{x})$ and, for each $n \geq 1$, $t \rightsquigarrow \mathbf{X}_n(-t, \mathbf{x})$ are given by the same prescription as $t \rightsquigarrow \mathbf{X}(t, \mathbf{x})$ and $t \rightsquigarrow \mathbf{X}_n(t, \mathbf{x})$ with $\mathbf{F}$ replaced by $-\mathbf{F}$. Hence, it suffices to handle $t \geq 0$.

To carry out this program, we will make frequent use of the following lemma, known as *Gromwall's inequality*.

Lemma 4.5.1 *Suppose that $\alpha : [0, \infty) \longrightarrow \mathbb{R}$ and $\beta : [0, \infty) \longrightarrow \mathbb{R}$ are continuous functions and that α is non-decreasing. If $T > 0$ and $u : [0, T] \longrightarrow \mathbb{R}$ is a continuous function that satisfies*

$$u(t) \leq \alpha(t) + \int_0^t \beta(\tau) u(\tau)\, d\tau \quad \textit{for } t \in [0, T],$$

then

$$u(T) \leq \alpha(0) e^{B(T)} + \int_0^T e^{B(T) - B(\tau)}\, d\alpha(\tau) \quad \textit{where } B(t) \equiv \int_0^t \beta(\tau)\, d\tau.$$

Proof Set $U(t) = \int_0^t \beta(\tau) u(\tau)\, d\tau$. Then $\dot{U}(t) \leq \alpha(t)\beta(t) + \beta(t)U(t)$, and so

$$\frac{d}{dt}\big(e^{-B(t)} U(t)\big) \leq \alpha(t)\beta(t) e^{-B(t)}.$$

Integrating both sides over $[0, T]$ and applying part (ii) of Exercise 3.16, we obtain

$$e^{-B(T)} U(t) \leq \int_0^T \alpha(\tau)\beta(\tau) e^{-B(\tau)}\, d\tau = -\alpha(T) e^{-B(T)} + \alpha(0) + \int_0^T e^{-B(\tau)}\, d\alpha(\tau).$$

Hence the required result follows from $u(T) \leq \alpha(T) + U(T)$. □

Our first application of Lemma 4.5.1 provides an estimate on the size of $\mathbf{X}_n$.

Lemma 4.5.2 *Assume that there exist $a \geq 0$ and $b > 0$ such that $|\mathbf{F}(\mathbf{x})| \leq a + b|\mathbf{x}|$ for all $\mathbf{x} \in \mathbb{R}^N$. If $\mathbf{X}_n(t, \mathbf{x})$ is given by (4.5.1), then*

$$\sup_{n \geq 1} |\mathbf{X}_n(t, \mathbf{x})| \leq |\mathbf{x}| e^{b|t|} + \frac{a\big(e^{b|t|} - 1\big)}{b} \quad \textit{for all } t \in \mathbb{R}.$$

Proof By our observation about $t \rightsquigarrow \mathbf{X}_n(-t, \mathbf{x})$, it suffices to handle $t \geq 0$.

Set $u_n(t) = \max\{|\mathbf{X}_n(\tau, \mathbf{x})| : t \in [0, t]\}$. Then

$$|\mathbf{X}_n(t, \mathbf{x})| \leq |\mathbf{x}| + \int_0^t \left|\mathbf{F}\big(\mathbf{X}_n\big(\lfloor\tau\rfloor_n, \mathbf{x}\big)\big)\right| d\tau \leq |\mathbf{x}| + at + b\int_0^t u_n(\tau)\,d\tau$$

and therefore

$$u_n(t) \leq |\mathbf{x}| + at + b\int_0^t u_n(\tau)\,d\tau \quad \text{for } t \geq 0.$$

Thus the desired estimate follows by Gromwall's inequality. □

From now on we will be assuming that $\mathbf{F}$ is globally *Lipschitz continuous*. That is, there exists an $L \in [0, \infty)$, called the *Lipschitz constant*, such that

$$|\mathbf{F}(\mathbf{y}) - \mathbf{F}(\mathbf{x})| \leq L|\mathbf{y} - \mathbf{x}| \quad \text{for all } \mathbf{x},\ \mathbf{y} \in \mathbb{R}^N. \tag{4.5.4}$$

By Schwarz's inequality, (4.5.4) will hold if $\mathbf{F}$ is continuously differentiable and

$$L = \sqrt{\sup_{\mathbf{x}\in\mathbb{R}^N} \sum_{j=1}^N |\nabla F_j(\mathbf{x})|^2} < \infty.$$

We will now show that (4.5.2) has at most one solution when (4.5.4) holds. To this end, suppose that $t \rightsquigarrow \mathbf{X}(t, \mathbf{x})$ and $t \rightsquigarrow \tilde{\mathbf{X}}(t, \mathbf{x})$ are two solutions, and set $u(t) = |\mathbf{X}(t, \mathbf{x}) - \tilde{\mathbf{X}}(t, \mathbf{x})|$. Then

$$u(t) \leq \int_0^t \left|\mathbf{F}\big(\mathbf{X}(\tau, \mathbf{x})\big) - \mathbf{F}\big(\tilde{\mathbf{X}}(\tau, \mathbf{x})\big)\right| d\tau \leq L\int_0^t u(\tau)\,d\tau \quad \text{for } t \geq 0,$$

and therefore by Gromwall's inequality with $\alpha \equiv 0$ and $\beta \equiv L$, we know that $u(t) = 0$ for all $t \geq 0$. To handle $t < 0$, use the observation about $t \rightsquigarrow \mathbf{X}(-t, \mathbf{x})$.

Theorem 4.5.3 *Assume that* $\mathbf{F}$ *satisfies (4.5.4). Then for each* $\mathbf{x} \in \mathbb{R}^N$ *there is a unique solution* $t \rightsquigarrow \mathbf{X}(t, \mathbf{x})$ *to (4.5.2). Furthermore, for each* $R > 0$ *there exists a* $C_R < \infty$*, depending only on the Lipschitz constant* L *in (4.5.4), such that*

$$\sup\big\{|\mathbf{X}(t, \mathbf{x}) - \mathbf{X}_n(t, \mathbf{x}) : |t| \vee |\mathbf{x}| \leq R\big\} \leq \frac{C_R}{n}. \tag{4.5.5}$$

Finally,

$$|\mathbf{X}(t, \mathbf{y}) - \mathbf{X}(t, \mathbf{x})| \leq e^{L|t|}|\mathbf{y} - \mathbf{x}| \quad \textit{for all } t \in \mathbb{R} \textit{ and } \mathbf{x},\ \mathbf{y} \in \mathbb{R}^N.$$

Proof The uniqueness assertion was proved above. To prove the existence of $t \rightsquigarrow \mathbf{X}(t, \mathbf{x})$, set $u_{m,n}(t, \mathbf{x}) = |\mathbf{X}_n(t, \mathbf{x}) - \mathbf{X}_m(t, \mathbf{x})|$ for $t \geq 0$ and $1 \leq m < n$. Then

$$\begin{aligned} u_{m,n}(t, \mathbf{x}) &\leq \int_0^t \left|\mathbf{F}\big(\mathbf{X}_n\big(\lfloor\tau\rfloor_n, \mathbf{x}\big)\big) - \mathbf{F}\big(\mathbf{X}_m\big(\lfloor\tau\rfloor_m, \mathbf{x}\big)\big)\right| d\tau \\ &\leq L \int_0^t \Big(\big|\mathbf{X}_n\big(\lfloor\tau\rfloor_n, \mathbf{x}\big) - \mathbf{X}_n(\tau, \mathbf{x})\big| + \big|\mathbf{X}_m\big(\lfloor\tau\rfloor_m, \mathbf{x}\big) - \mathbf{X}_m(\tau, \mathbf{x})\big|\Big)\, d\tau \\ &\qquad + L \int_0^t u_{m,n}(\tau, \mathbf{x})\, d\tau. \end{aligned}$$

Next observe that $|\mathbf{F}(\mathbf{x})| \leq |\mathbf{F}(0)| + L|\mathbf{x}|$, and therefore, by Lemma 4.5.2,

$$|\mathbf{X}_k(t, \mathbf{x})| \leq Re^{LR} + \frac{|\mathbf{F}(\mathbf{0})|\big(e^{LR} - 1\big)}{L} \quad \text{for } k \geq 1 \text{ and } (t, \mathbf{x}) \in [0, R] \times \overline{B(\mathbf{0}, R)}.$$

Thus, if $A_R = |\mathbf{F}(0)| + \big(LR + |\mathbf{F}(\mathbf{0})|(1 - e^{-LR})\big)e^{LR}$, then

$$\left|\mathbf{X}_k\big(\lfloor\tau\rfloor_k, \mathbf{x}\big) - \mathbf{X}_k(\tau, \mathbf{x})\right| \leq \int_{\lfloor\tau\rfloor_k}^t \left|\mathbf{F}\big(\mathbf{X}_k(\sigma, \mathbf{x})\big)\right| d\sigma \leq \frac{A_R}{k}$$

for $k \geq 1$ and $(\tau, \mathbf{x}) \in [0, R] \times \overline{B(\mathbf{0}, R)}$. Hence, we now know that, if $(t, \mathbf{x}) \in [0, T] \times \overline{B(\mathbf{0}, R)}$, then

$$u_{m,n}(t, \mathbf{x}) \leq \frac{2A_R}{m} + L \int_0^t u_{m,n}(\tau, \mathbf{x})\, d\tau,$$

and therefore, by Gromwall's inequality,

$$(*) \qquad \sup_{(t,\mathbf{x}) \in [0,R] \times \overline{B(\mathbf{0},R)}} |\mathbf{X}_n(t, \mathbf{x}) - \mathbf{X}_m(t, \mathbf{x})| \leq \frac{C_R}{m},$$

where $C_R = 2A_R e^{LR}$.

From $(*)$ it is clear that, for each $t \geq 0$, $\{\mathbf{X}_n(t, \mathbf{x}) : n \geq 1\}$ satisfies Cauchy's convergence criterion and therefore converges to some $\mathbf{X}(t, \mathbf{x})$. Further, because, again from $(*)$,

$$\sup_{(t,\mathbf{x}) \in [0,R] \times \overline{B(\mathbf{0},R)}} |\mathbf{X}(t, \mathbf{x}) - \mathbf{X}_m(t, \mathbf{x})| \leq \frac{C_R}{m},$$

(4.5.5) holds with this C_R. Hence, by Lemma 1.4.4, $t \rightsquigarrow \mathbf{X}(t, \mathbf{x})$ is continuous, and by Theorem 3.1.4, (4.5.2) follows from (4.5.1) for $t \geq 0$. To prove the same result for $t < 0$, one again uses the observation made earlier.

Finally, observe that

$$|\mathbf{X}(t,\mathbf{y}) - \mathbf{X}(t,\mathbf{x})| \le |\mathbf{y}-\mathbf{x}| + L\int_0^{|t|} |\mathbf{X}(\tau,\mathbf{y}) - \mathbf{X}(\tau,\mathbf{x})|\,d\tau \quad \text{for } t \in \mathbb{R},$$

and now proceed as before, via Gromwall's inequality, to get the estimate in the concluding assertion. □

Theorem 4.5.3 is a basic result that guarantees the existence and uniqueness of solutions to the ordinary differential equation (4.5.3). It should be pointed out that the existence result holds under much weaker assumptions (cf. Exercise 4.19). For example, solutions will exist if $\mathbf{F}$ is continuous and $|\mathbf{F}(x)| \le a + b|\mathbf{x}|$ for some non-negatives constants a and b. On the other hand, uniqueness can fail when $\mathbf{F}$ is not Lipschitz continuous. To wit, consider the equation $\dot{X}(t) = |X(t)|^{\frac{1}{2}}$ with $X(0) = 0$. Obviously, one solution is $X(t) = 0$ for all $t \in \mathbb{R}$. On the other hand, a second solution is

$$X(t) = \begin{cases} \frac{t^2}{4} & \text{if } t \ge 0 \\ -\frac{t^2}{4} & \text{if } t < 0. \end{cases}$$

For many purposes, the uniqueness is just as important as existence. In particular, it allows one to prove that

$$\mathbf{X}(s+t,\mathbf{x}) = \mathbf{X}\big(t,\mathbf{X}(s,\mathbf{x})\big) \quad \text{for all } s,\ t \in \mathbb{R} \text{ and } \mathbf{x} \in \mathbb{R}^N. \tag{4.5.6}$$

Indeed, observe that $t \rightsquigarrow \mathbf{Y}(t) \equiv \mathbf{X}(s+t,\mathbf{x})$ satisfies $\dot{\mathbf{Y}}(t) = \mathbf{F}\big(\mathbf{Y}(t)\big)$ with $\mathbf{Y}(0) = \mathbf{X}(s,\mathbf{x})$, and therefore $\mathbf{Y}(t) = \mathbf{X}\big(t,\mathbf{X}(s,\mathbf{x})\big)$. The equality in (4.5.6) is called the *flow property* and reflects the fact that once one knows where a solution to (4.5.3) is at a time s its position at any other time is completely determined. An important consequence of (4.5.6) is that, for any $t \in \mathbb{R}$, $\mathbf{y} \rightsquigarrow \mathbf{X}(-t,\mathbf{y})$ is the inverse of $\mathbf{x} \rightsquigarrow \mathbf{X}(t,\mathbf{x})$ and therefore, when $\mathbf{F}$ satisfies (4.5.4), that $\mathbf{x} \rightsquigarrow \mathbf{X}(t,\mathbf{x})$ and its inverse are continuous, one-to-one maps of $\mathbb{R}^N$ onto itself.

Corollary 4.5.4 *Assume that* $\mathbf{F} : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ *is continuously differentiable and that* $\partial_{\mathbf{e}_j}\mathbf{F}$ *is bounded for each* $1 \le j \le N$*. Then, for each* $t \in \mathbb{R}$*,* $\mathbf{x} \rightsquigarrow \mathbf{X}(t,\mathbf{x})$ *is continuously differentiable and*

$$\partial_{\mathbf{e}_j}\mathbf{X}(t,\mathbf{x})_i = \delta_{i,j} + \int_0^t \Big(\nabla F_i\big(\mathbf{X}(\tau,\mathbf{x})\big), \partial_{\mathbf{e}_j}\mathbf{X}(\tau,\mathbf{x})\Big)_{\mathbb{R}^N}\,d\tau. \tag{4.5.7}$$

Proof Since, as we have seen, results for $t \le 0$ can be obtained from the ones for $t \ge 0$, we will restrict our attention to $t \ge 0$

First observe that, for each $n \ge 1$ and $t \ge 0$, $\mathbf{x} \rightsquigarrow \mathbf{X}_n(t,\mathbf{x})$ is continuously differentiable and satisfies (cf. Exercise 4.7)

$$\partial_{\mathbf{e}_j}\mathbf{X}_n(t,\mathbf{x})_i = \delta_{i,j} + \int_0^t \Big(\nabla F_i\big(\mathbf{X}_n\big(\lfloor\tau\rfloor_n,\mathbf{x}\big), \partial_{\mathbf{e}_j}\mathbf{X}_n\big(\lfloor\tau\rfloor_n,\mathbf{x}\big)\Big)_{\mathbb{R}^N}\,d\tau.$$

Thus

$$|\partial_{\mathbf{e}_j}\mathbf{X}_n(t,\mathbf{x})| \le 1 + A\int_0^t |\partial_{\mathbf{e}_j}\mathbf{X}_n(\lfloor\tau\rfloor_n,\mathbf{x})|\,d\tau$$

for $n \ge 1$ and $(t,\mathbf{x}) \in [0,\infty)\times\mathbb{R}^N$, where $A = \|\,|\nabla F_i|\,\|_{\mathbb{R}^N} < \infty$. Hence, arguing as we did in the proof of Lemma 4.5.2, we see that $|\partial_{\mathbf{e}_j}\mathbf{X}_n(t,\mathbf{x})| \le e^{At}$, and therefore that there exists a $B < \infty$ such that

$$\left|\partial_{\mathbf{e}_j}\mathbf{X}_n(t,\mathbf{x}) - \partial_{\mathbf{e}_j}\mathbf{X}_n(s,\mathbf{x})\right| \le Be^{At}|t-s|$$

for $n \ge 1$ and $0 \le s < t$ and $\mathbf{x} \in \mathbb{R}^N$. Combining this with (4.5.5) and the fact the first derivatives of $\mathbf{F}$ are continuous, we conclude that

$$\sup_{n>m}\sup_{t\vee|\mathbf{x}|\le R}\max_{1\le i\le N}\left|\nabla F_i\big(\mathbf{X}_n(\lfloor t\rfloor_n,\mathbf{x})\big) - \nabla F_i\big(\mathbf{X}_m(\lfloor t\rfloor_m,\mathbf{x})\big)\right|$$

tends to 0 as $m \to \infty$. Hence there exist $\{\epsilon_m(R) : m \ge 1\} \subseteq (0,\infty)$ such that $\epsilon_m(R) \searrow 0$ as $m \to \infty$ and

$$\left|\partial_{\mathbf{e}_j}\mathbf{X}_n(t,\mathbf{x}) - \partial_{\mathbf{e}_j}\mathbf{X}_m(t,\mathbf{x})\right| \le \epsilon_m(R) + AN^{\frac{1}{2}}\int_0^t \left|\partial_{\mathbf{e}_j}\mathbf{X}_n(\tau,\mathbf{x}) - \partial_{\mathbf{e}_j}\mathbf{X}_m(\tau,\mathbf{x})\right|d\tau$$

for $1 \le m < n$ and $t \vee |\mathbf{x}| \le R$. Therefore, by Gromwall's inequality, we now know that

$$\sup_{n>m}\left|\partial_{\mathbf{e}_j}\mathbf{X}_n(t,\mathbf{x}) - \partial_{\mathbf{e}_j}\mathbf{X}_m(t,\mathbf{x})\right| \longrightarrow 0$$

uniformly for $t \vee |\mathbf{x}| \le R$.

From here, the rest of the argument is easy. By Cauchy's criterion and Lemma 1.4.4, there exists a continuous $\mathbf{Y}_j : [0,\infty)\times\mathbb{R}^N \longrightarrow \mathbb{R}^N$ to which $\partial_{\mathbf{e}_j}\mathbf{X}_n(t,\mathbf{x})$ converges uniformly on bounded subsets of $[0,\infty)\times\mathbb{R}^N$, and so, by Corollary 3.2.4, $\mathbf{x} \rightsquigarrow \mathbf{X}(t,\mathbf{x})$ is continuously differentiable and $\partial_{\mathbf{e}_j}\mathbf{X}(t,\mathbf{x}) = \mathbf{Y}_j(t,\mathbf{x})$. Furthermore, by Theorem 3.1.4, $t \rightsquigarrow \partial_{\mathbf{e}_j}\mathbf{X}(t,\mathbf{x})$ satisfies (4.5.7). □

The following corollary gives one of the reasons why it is important to have these results for vector fields on $\mathbb{R}^N$ instead of just functions.

Corollary 4.5.5 *Suppose that $F : \mathbb{R}^N \longrightarrow \mathbb{R}$ satisfies (4.5.4) for some $L < \infty$. Then for each $\mathbf{x} \in \mathbb{R}^N$ there is precisely one function $t \in \mathbb{R} \longmapsto X(t,\mathbf{x}) \in \mathbb{R}$ such that*

$$\begin{aligned}\partial_t^N X(t,x) = F\big(X(t,\mathbf{x}),\partial_t^1 X(t,\mathbf{x}),\dots,\partial_t^{N-1}X(t,x)\big)\\ \text{with } \partial_t^k X(0,\mathbf{x}) = x_{k+1} \text{ for } 0 \le k < N,\end{aligned} \tag{4.5.8}$$

where $\partial_t^k X(t,\mathbf{x})$ denotes the kth derivative of $X(\,\cdot\,,\mathbf{x})$ with respect to t.

Proof Define $\mathbf{F}(\mathbf{x}) = \big(x_2, \dots, x_N, F(\mathbf{x})\big)$ for $\mathbf{x} \in \mathbb{R}^N$. Then $\mathbf{F}$ satisfies (4.5.4), and therefore there is precisely one solution to

$$(*) \qquad \dot{\mathbf{X}}(t,\mathbf{x}) = \mathbf{F}\big(\mathbf{X}(t,\mathbf{x})\big) \quad \text{with } \mathbf{X}(0,\mathbf{x}) = \mathbf{x}.$$

Furthermore, $t \rightsquigarrow \mathbf{X}(t,\mathbf{x})$ satisfies $(*)$ if and only if $\dot{X}_k(t,\mathbf{x}) = X_{k+1}(t,\mathbf{x})$ for $0 \le k < N$, $\dot{X}_N(t,x) = F\big(\mathbf{X}(t,\mathbf{x})\big)$, and $X_k(0,\mathbf{x}) = x_{k+1}$ for $1 \le k < N$. Hence, if $X(t,\mathbf{x}) \equiv X_1(t,\mathbf{x})$, then $\partial_t^k X(t,\mathbf{x}) = X_{k+1}(t,\mathbf{x})$ for $0 \le k < N$ and

$$\partial_t^N X(t,\mathbf{x}) = F\big(X(t,\mathbf{x}), \partial_t^1 X(t,\mathbf{x}), \dots, \partial_t^{N-1} X(t,\mathbf{x})\big),$$

which means that $t \rightsquigarrow X(t,\mathbf{x})$ is a solution to (4.5.8). Conversely, if $t \rightsquigarrow X(t,\mathbf{x})$ is a solution to (4.5.8) and

$$\mathbf{X}(t,\mathbf{x}) = \big(X(t,\mathbf{x}), \partial_t^1 X(t,\mathbf{x}), \dots, \partial_t^{N-1} X(t,\mathbf{x})\big),$$

then $t \rightsquigarrow \mathbf{X}(t,\mathbf{x})$ is a solution to $(*)$, and so $t \rightsquigarrow X(t,\mathbf{x})$ is the only solution to (4.5.8). □

The idea on which the preceding proof is based was introduced by physicists to study Newton's equation of motion, which is a second order equation. What they realized is that the analysis of his equation is simplified if one moves to what they called phase space, in which points represent the position and velocity of a particle and Newton's equation becomes a first order equation.

4.6 Exercises

Exercise 4.1 Given a set $S \subseteq \mathbb{R}^N$, the set $\partial S \equiv \bar{S} \setminus \text{int}(S)$ is called the *boundary* of S. Show that $\mathbf{x} \in \partial S$ if and only if $B(\mathbf{x}, r) \cap S \ne \emptyset$ and $B(\mathbf{x}, r) \cap S\complement \ne \emptyset$ for every $r > 0$. Use this characterization to show that $\partial S = \partial(S\complement)$, $\partial(S_1 \cup S_2) \subseteq \partial S_1 \cup \partial S_2$, $\partial(S_1 \cap S_2) \subseteq \partial S_1 \cup \partial S_2$, and $\partial(S_2 \setminus S_1) \subseteq \partial S_1 \cup \partial S_2$.

Exercise 4.2 Suppose that $\{\mathbf{x}_n : n \ge 1\}$ is a bounded sequence in $\mathbb{R}^N$. Show that $\{\mathbf{x}_n : n \ge 1\}$ converges if and only if it has only one limit point. That is, $\mathbf{x}_n \longrightarrow \mathbf{x}$ if and only if every convergent subsequence $\{\mathbf{x}_{n_k} : k \ge 1\}$ converges to $\mathbf{x}$.

Exercise 4.3 Let F_1 and F_2 be a pair of disjoint, closed subsets of $\mathbb{R}^N$. Assuming that at least one of them is bounded, show that there exist $\mathbf{x}_1 \in F_1$ and $\mathbf{x}_2 \in F_2$ such that

$$|\mathbf{x}_2 - \mathbf{x}_1| = |F_2 - F_1| \equiv \inf\{|\mathbf{y}_2 - \mathbf{y}_1| : \mathbf{y}_1 \in F_1 \ \&\ \mathbf{y}_2 \in F_2\},$$

and conclude that $|F_2 - F_1| > 0$. On the other hand, give an example of disjoint, closed subsets F_1 and F_2 of $\mathbb{R}$ such that $|F_1 - F_1| = 0$.

Exercise 4.4 Given a pair of closed sets F_1 and F_2, set

$$F_1 + F_2 = \{\mathbf{x}_1 + \mathbf{x}_2 : \mathbf{x}_1 \in F_1 \ \& \ \mathbf{x}_2 \in F_2\}.$$

Show that $F_1 + F_2$ is closed if at least one of the sets F_1 and F_2 is bounded but that it need not be closed if neither is bounded.

Exercise 4.5 Given a non-empty open subset G of $\mathbb{R}^N$ and a point $\mathbf{x} \in G$, let $G_\mathbf{x}$ be the set of $\mathbf{y} \in G$ to which $\mathbf{x}$ is connected in the sense that there exists a continuous path $\gamma : [a, b] \longrightarrow G$ such that $\mathbf{x} = \gamma(a)$ and $\mathbf{y} = \gamma(b)$. Show that both $G_\mathbf{x}$ and $G \setminus G_\mathbf{x}$ are open. Next, say that G is *connected* if it cannot be written as the union of two non-empty open sets, and show that G is connected if and only if $G_\mathbf{x} = G$ for every $\mathbf{x} \in G$. Thus G *is connected if and only if it path connected.*

Exercise 4.6 Suppose that K is a compact subset of $\mathbb{R}^N$ and that $\mathbf{F}$ is an $\mathbb{R}^M$-valued continuous function on K. Show that $\mathbf{F}(K)$ is a compact subset of $\mathbb{R}^M$. The fact that $\mathbf{F}(K)$ is bounded is obvious, what is interesting is that it is closed. Indeed, give an example of a continuous function $f : \mathbb{R} \longrightarrow \mathbb{R}$ for which $f(\mathbb{R})$ is not closed.

Exercise 4.7 Suppose that G_1 and G_2 are non-empty, open subsets of $\mathbb{R}^{N_1}$ and $\mathbb{R}^{N_2}$ respectively. Further, assume that $F_1, \dots, F_{N_2}$ are continuously differentiable functions on G_1 and that $\mathbf{F}(\mathbf{x}) \equiv \big(F_1(\mathbf{x}), \dots, F_{N_2}(\mathbf{x})\big) \in G_2$ for all $\mathbf{x} \in G_1$. Finally, let $g : G_2 \longrightarrow \mathbb{C}$ be continuously differentiable, and define $g \circ \mathbf{F}(\mathbf{x}) = g\big(\mathbf{F}(\mathbf{x})\big)$ for $\mathbf{x} \in G_1$. Show that

$$\partial_{\boldsymbol{\xi}}(g \circ \mathbf{F})(\mathbf{x}) = \sum_{j=1}^{N_2} (\partial_{\mathbf{e}_j} g) \circ \mathbf{F}(\mathbf{x}) \, \partial_{\boldsymbol{\xi}} F_j(\mathbf{x})$$

for $\mathbf{x} \in G_1$ and $\boldsymbol{\xi} \in \mathbb{R}^{N_1}$. This is the form that the *chain rule* takes for functions of several variables.

Exercise 4.8 Let $\mathcal{G}$ be a collection of open subsets of $\mathbb{R}^N$. The goal of this exercise is to show that there exists a sequence $\{G_k : k \geq 1\} \subseteq \mathcal{G}$ such that $\bigcup_{k=1}^\infty G_k = \bigcup_{G \in \mathcal{G}} G$. This is sometimes called the *Lindelöf property* of $\mathbb{R}^N$. Here are some steps that you might take.

(i) Let $\mathcal{A}$ be the set of pairs $(\mathbf{q}, \ell)$ where $\mathbf{q}$ is an element of $\mathbb{R}^N$ with rational coordinates and $\ell \in \mathbb{Z}^+$. Show that $\mathcal{A}$ is countable.

(ii) Given an open set G, let $\mathcal{B}_G$ be the set of balls $B\big(\mathbf{q}, \frac{1}{\ell}\big)$ such that $(\mathbf{q}, \ell) \in \mathcal{A}$ and $B\big(\mathbf{q}, \frac{1}{\ell}\big) \subseteq G$. Show that $G = \bigcup_{B \in \mathcal{B}_G} B$.

(iii) Given a collection $\mathcal{G}$ of open sets, let $\mathcal{B}_\mathcal{G} = \bigcup_{G \in \mathcal{G}} \mathcal{B}_G$. Show that $\mathcal{B}_\mathcal{G}$ is countable and that $\bigcup_{B \in \mathcal{B}_\mathcal{G}} B = \bigcup_{G \in \mathcal{G}} G$. Finally, for each $B \in \mathcal{B}_\mathcal{G}$, choose a $G_B \in \mathcal{G}$ so that $B \subseteq G_B$, and observe that $\bigcup_{B \in \mathcal{B}_\mathcal{G}} G_B = \bigcup_{G \in \mathcal{G}} G$.

Exercise 4.9 According to Lemma 4.1.1, a set $K \subseteq \mathbb{R}^N$ is compact if and only if it is closed and bounded. In this exercise you are to show that K is compact if and only if for each collection of open sets $\mathcal{G}$ that covers K (i.e., $K \subseteq \bigcup_{G\in\mathcal{G}} G$) there is a finite sub-cover $\{G_1, \dots, G_n\} \subseteq \mathcal{G}$ which covers K (i.e., $K \subseteq \bigcup_{m=1}^n G_m$). Here is one way that you might proceed.

(i) Assume that every cover $\mathcal{G}$ of K by open sets admits a finite sub-cover. To show that K is bounded, take $\mathcal{G} = \{B(\mathbf{x}, 1) : \mathbf{x} \in K\}$ and pass to a finite sub-cover. To see that K is closed, suppose that $\mathbf{y} \notin K$, and, for $m \geq 1$, define G_m to be the set of $\mathbf{x} \in \mathbb{R}^N$ such that $|\mathbf{y} - \mathbf{x}| > \frac{1}{m}$. Show that each G_m is open and that $K \subseteq \bigcup_{m=1}^\infty G_m$, and choose $n \geq 1$ so that $K \subseteq \bigcup_{m=1}^n G_m$. Show that $B\big(\mathbf{y}, \frac{1}{n}\big) \cap K = \emptyset$ and therefore that $\mathbf{y} \notin \bar{K}$. Hence $K = \bar{K}$.

(ii) Assume that K is compact, and let $\mathcal{G}$ be a cover of K by open sets. Using Exercise 4.8, choose $\{G_m : m \geq 1\} \subseteq \mathcal{G}$ so that $K \subseteq \bigcup_{m=1}^\infty G_m$. Suppose that $K \nsubseteq \bigcup_{m=1}^n G_m$ for any $n \geq 1$, and, for each $n \geq 1$, choose $\mathbf{x}_n \in K$ so that $\mathbf{x}_n \notin \bigcup_{m=1}^n G_m$. Now let $\{\mathbf{x}_{n_j} : j \geq 1\}$ be a subsequence that converges to a point $\mathbf{x} \in K$, and choose $m \geq 1$ such that $\mathbf{x} \in G_m$. Then there would exist a $j \geq 1$ such that $n_j \geq m$ and $\mathbf{x}_{n_j} \in G_m$, which is impossible.

Exercise 4.10 Here is a typical example that shows how the result in Exercise 4.9 gets applied. Given a set $S \subseteq \mathbb{R}^N$ and a family $\mathcal{F}$ of functions $f : S \longrightarrow \mathbb{C}$, one says that $\mathcal{F}$ is *equicontinuous* at $\mathbf{x} \in S$ if, for each $\epsilon > 0$, there is a $\delta > 0$ such that $|f(\mathbf{y}) - f(\mathbf{x})| < \epsilon$ for all $\mathbf{y} \in S \cap B(\mathbf{x}, \delta)$ and all $f \in \mathcal{F}$. Show that if $\mathcal{F}$ is equicontinuous at each $\mathbf{x}$ in a compact set K, then it is *uniformly equicontinuous* in the sense that for each $\epsilon > 0$ there is a $\delta > 0$ such that $|f(\mathbf{y}) - f(\mathbf{x})| < \epsilon$ for all $\mathbf{x}, \mathbf{y} \in K$ with $|\mathbf{y} - \mathbf{x}| < \delta$ and all $f \in \mathcal{F}$. The idea is to begin by choosing for each $\mathbf{x} \in K$ a $\delta_{\mathbf{x}} > 0$ such that $\sup_{f\in\mathcal{F}} |f(\mathbf{y}) - f(\mathbf{x})| < \frac{\epsilon}{2}$ for all $\mathbf{y} \in K \cap B(\mathbf{x}, 2\delta_x)$. Next, choose $\mathbf{x}_1, \dots, \mathbf{x}_n \in K$ so that $K \subseteq \bigcup_{m=1}^n B(\mathbf{x}_m, \delta_{\mathbf{x}_m})$. Finally, check that $\delta = \min\{\delta_{\mathbf{x}_m} : 1 \leq m \leq n\}$ has the required property.

Exercise 4.11 Another application of Exercise 4.9 is provided by *Dini's lemma*. Suppose that $\{f_n : n \geq 1\}$ is a non-increasing sequence of $\mathbb{R}$-valued, continuous functions on a compact set K, and assume that f is a continuous function of K to which they converge pointwise. Show that $f_n \longrightarrow f$ uniformly on K. In doing this problem, first reduce to the case when $f = 0$. Then, given $\epsilon > 0$, for each $\mathbf{x} \in K$ choose $n_{\mathbf{x}} \in \mathbb{Z}^+$ and $r_{\mathbf{x}} > 0$ so that $f_{n_{\mathbf{x}}}(\mathbf{y}) \leq \epsilon$ for $\mathbf{y} \in B(\mathbf{x}, r_{\mathbf{x}})$. Now apply Exercise 4.9 to the cover $\{B(\mathbf{x}, r_{\mathbf{x}}) : \mathbf{x} \in K\}$.

Exercise 4.12 In this exercise you are to construct a *space-filling curve*. That is, for each $N \geq 2$, a continuous map $\mathbf{X}$ that takes the interval $[0, 1]$ *onto* the square $[0, 1]^N$. Such a curve was constructed originally in the 19th Century by G. Peano, but the one suggested here is much simpler than Peano's and was introduced by I. Scheonberg. Let $f : \mathbb{R} \longrightarrow [0, 1]$ be any continuous function with the properties that $f(t) = 0$ if $\big[0, \frac{1}{3}\big]$, $f(t) = 1$ if $t \in \big[\frac{2}{3}, 1\big]$, $f(2) = 0$, and $f(t + 2) = f(t)$ for all $t \in \mathbb{R}$. For example, one can take $f(t) = 0$ for $t \in [0, \frac{1}{3}]$, $f(t) = 3\big(t - \frac{1}{3}\big)$ for $t \in \big[\frac{1}{3}, \frac{2}{3}\big]$, $f(t) = 1$ for $t \in \big[\frac{2}{3}, 1\big]$, $f(t) = 2 - t$ for $t \in [1, 2]$, and then define $f(t) = f(t - 2n)$ for $n \in \mathbb{Z}$ and $2n \leq t < 2(n + 1)$. Next, define $\mathbf{X}(t) = \big(X_1(t), \dots, X_N(t)\big)$ where

$$X_j(t) = \sum_{n=0}^{\infty} 2^{-n-1} f\big(3^{nN+j-1}\big) \quad \text{for } 1 \le j \le N.$$

Given $\mathbf{x} = (x_1, \dots, x_N) \in [0, 1]^N$, choose $\omega : \mathbb{N} \longrightarrow \{0, 1\}$ so that

$$x_j = \sum_{n=0}^{\infty} 2^{-n-1} \omega(nN + j - 1) \quad \text{for } 1 \le j \le N,$$

set $s = 2\sum_{n=0}^{\infty} 3^{-n-1}\omega(n)$, and show that $\mathbf{X}(s) = \mathbf{x}$. To this end, observe that, for each $m \in \mathbb{N}$, there is a $b_m \in \mathbb{N}$ such that

$$2b_m + \frac{2\omega(m)}{3} \le 3^m s \le 2b_m + \frac{2\omega(m)}{3} + \frac{1}{3}$$

and therefore that $f(3^m s) = \omega(m)$.

Obviously, $\mathbf{X}$ is very far from being one-to-one, and there are good topological reasons why it cannot be. Finally, recall the Cantor set C in Exercise 1.8, and show that the restriction of $\mathbf{X}$ to C is a one-to-one map onto $[0, 1]^N$. This, of course, gives another proof that C is uncountable.

Exercise 4.13 Define $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ so that, for $\mathbf{x} = (x_1, x_2)$, $f(\mathbf{x}) = \frac{x_1^2 x_2^2}{x_1^4 + x_2^4}$ if $\mathbf{x} \ne \mathbf{0}$ and $f(\mathbf{x}) = 0$ if $\mathbf{x} = \mathbf{0}$. Show that f is differentiable at $\mathbf{0}$ and that $\partial_{\boldsymbol{\xi}} f(\mathbf{0}) = f(\boldsymbol{\xi})$. Conclude in particular that for this f it is *not* true that $\partial_{\boldsymbol{\xi}} f(\mathbf{0}) = (\boldsymbol{\xi}, \mathbf{e}_1)_{\mathbb{R}^2} \partial_{\mathbf{e}_1} f(\mathbf{0}) + (\boldsymbol{\xi}, \mathbf{e}_2)_{\mathbb{R}^2} \partial_{\mathbf{e}_2} f(\mathbf{0})$ for all $\boldsymbol{\xi}$.

Exercise 4.14 Let $a, b, c \in \mathbb{R}$, and show that

$$a\xi^2 + 2b\xi\eta + c\eta^2 \ge 0 \text{ for all } \xi, \eta \in \mathbb{R} \iff a + c \ge 0 \text{ and } ac \ge b^2.$$

Now let $f : G \longrightarrow \mathbb{R}$ be a twice continuously differentiable function on a non-empty open set $G \subseteq \mathbb{R}^2$, and, given $\mathbf{x} \in G$, show that $\partial_{\boldsymbol{\xi}}^2 f(\mathbf{x}) = (\boldsymbol{\xi}, Hf(\mathbf{x})\boldsymbol{\xi})_{\mathbb{R}^2} \ge 0$ for all $\boldsymbol{\xi} \in \mathbb{R}^2$ if and only if

$$\partial_{\mathbf{e}_1}^2 f(\mathbf{x}) + \partial_{\mathbf{e}_2}^2 f(\mathbf{x}) \ge 0 \text{ and } \partial_{\mathbf{e}_1}^2 f(\mathbf{x}) \partial_{\mathbf{e}_2}^2 f(\mathbf{x}) \ge \big(\partial_{\mathbf{e}_1}\partial_{\mathbf{e}_2} f(\mathbf{x})\big)^2.$$

Exercise 4.15 A set $C \subseteq \mathbb{R}^N$ is said to be *convex* if $(1-\theta)\mathbf{x} + \theta\mathbf{y} \in C$ whenever $\mathbf{x}, \mathbf{y} \in C$ and $\theta \in [0, 1]$. In other words, C contains the line segment connecting any pair of points in C. Show that $\bar{C}$ is convex if C is. Next, if f is an $\mathbb{R}$-valued function on a convex set C, f is called a *convex function* if $f\big((1-\theta)\mathbf{x} + \theta\mathbf{y}\big) \le (1-\theta)f(\mathbf{x}) + \theta f(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in C$ and $\theta \in [0, 1]$. Now let G be a non-empty open, convex set, set $C = \bar{G}$, and assume that $f : C \longrightarrow \mathbb{R}$ is continuous. If f is differentiable on G, show that f is convex on C if and only if for each $\mathbf{x} \in G$ and $r > 0$ such that $B(\mathbf{x}, r) \subseteq G$, $t \in [0, r) \longmapsto \partial_{\mathbf{e}} f(\mathbf{x} + t\mathbf{e}) \in \mathbb{R}$ is non-decreasing for all $\mathbf{e} \in \mathbb{R}^N$ with $|\mathbf{e}| = 1$. In particular, conclude that if f is twice continuously differentiable on G, then f is convex on C if and only if $\big(\boldsymbol{\xi}, Hf(\mathbf{x})\boldsymbol{\xi}\big)_{\mathbb{R}^N} \ge 0$ for all $\mathbf{x} \in G$ and $\boldsymbol{\xi} \in \mathbb{R}^N$.

Exercise 4.16 Prove the $\mathbb{R}^N$-version of (3.2.2). That is, if f is a $\mathbb{C}$-valued function on an open set $G \subseteq \mathbb{R}^N$ is $(n+1)$-times continuously differentiable and $(1-t)\mathbf{x} + t\mathbf{y} \in G$ for $t \in [0, 1]$, show that

$$f(\mathbf{y}) = \sum_{\|\mathbf{m}\| \le n} \frac{\partial^{\mathbf{m}} f(\mathbf{x})}{\mathbf{m}!}(\mathbf{y}-\mathbf{x})^{\mathbf{m}} + \frac{1}{n!}\int_0^1 (1-t)^n \partial_t^{n+1} f\big((1-t)\mathbf{x} + t\mathbf{y}\big)\,dt.$$

Exercise 4.17 Consider equations of the sort in Corollary 4.5.5. Even though one is looking for $\mathbb{R}$-valued solutions, it is sometimes useful to begin by looking for $\mathbb{C}$-valued ones. For example, consider the equation

$$(*) \qquad \partial_t^N Z(t) = \sum_{n=0}^{N-1} a_n \partial_t^n Z(t),$$

where $a_0, \dots, a_{N-1} \in \mathbb{R}$.

(i) Show that if Z is a $\mathbb{C}$-valued solution to $(*)$, then so is $\bar{Z}$. In addition, show that if Z_1 and Z_2 are two $\mathbb{C}$-valued solutions, then, for all $c_1,\ c_2 \in \mathbb{C}$, $c_1 Z_1 + c_2 Z_2$ is again a $\mathbb{C}$-valued solution. Conclude that if Z is a $\mathbb{C}$-valued solution, then both the real and the imaginary parts of Z are $\mathbb{R}$-valued solutions.

(ii) Set $P(z) = z^N - \sum_{n=0}^{N-1} a_n z^n$ for $z \in \mathbb{C}$. Show that

$$(**) \qquad \left(\partial_t^N - \sum_{n=0}^{N-1} \partial_t^n\right) e^{zt} = P(z)e^{zt}.$$

Next, assume that $\lambda \in \mathbb{C}$ is an ℓth order root of P (i.e., $P(z) = (z-\lambda)^\ell Q(z)$ for some polynomial Q) for some $1 \le \ell < N$, and use $(**)$ to show that $t^k e^{\lambda t}$ and $t^k e^{\bar{\lambda} t}$ are solutions for each $0 \le k < \ell$. In particular, if $\lambda = \alpha + i\beta$ where $\alpha,\ \beta \in \mathbb{R}$, then both $t^k e^{\alpha t} \cos \beta t$ and $t^k e^{\alpha t} \sin \beta t$ are $\mathbb{R}$-valued solutions for each $0 \le k < \ell$.

(iii) Consider the equation

$$\ddot{X}(t) = a\dot{X}(t) + bX(t) \quad \text{where } a,\ b \in \mathbb{R},$$

and set $D = a^2 + 4b$. Define

$$X_0(t) = e^{\frac{at}{2}} \times \begin{cases} \cosh\left(\sqrt{\frac{D}{2}}t\right) - a\sqrt{\frac{1}{2D}}\sinh\left(\sqrt{\frac{D}{2}}t\right) & \text{if } D > 0 \\ 1 - \frac{at}{2} & \text{if } D = 0 \\ \cos\left(\sqrt{\frac{|D|}{2}}t\right) - a\sqrt{\frac{1}{2|D|}}\sin\left(\sqrt{\frac{|D|}{2}}t\right) & \text{if } D < 0 \end{cases}$$

and

$$X_1(t) = e^{\frac{at}{2}} \times \begin{cases} \sqrt{\frac{2}{D}} \sinh\left(\sqrt{\frac{D}{2}} t\right) & \text{if } D > 0 \\ t & \text{if } D = 0 \\ \sqrt{\frac{1}{2|D|}} \sin\left(\sqrt{\frac{|D|}{2}} t\right) & \text{if } D < 0, \end{cases}$$

and show that $X : \mathbb{R} \longrightarrow \mathbb{R}$ is a solution if and only if

$$X(t) = X(0)X_0(t) + \dot{X}(0)X_1(t).$$

Exercise 4.18 When the condition in Lemma 4.5.2 fails to hold, solutions to (4.5.2) may "explode". For example, suppose that $\dot{X}(t) = X(t)|X(t)|^\alpha$ for some $\alpha > 0$. If $X(0) = x > 0$, show that $X(t) = (x - \alpha t)^{-\frac{1}{\alpha}}$ for $t \in \left(-\infty, \frac{x}{\alpha}\right)$ and therefore that $X(t)$ tends to ∞ as $t \nearrow \frac{x}{\alpha}$.

Exercise 4.19 As was already mentioned, solutions to (4.5.3) exist under a much weaker condition than that in (4.5.4), although uniqueness may not hold. Here we will examine a case in which both existence and uniqueness hold in the absence of (4.5.4). Namely, let $F : \mathbb{R} \longrightarrow (0, \infty)$ be a continuous function with the property that

$$\int_{(-\infty,0]} \frac{1}{F(s)}\, ds = \int_{[0,\infty)} \frac{1}{F(s)}\, ds = \infty,$$

and set

$$T_x(s) = \int_0^s \frac{1}{F(x+\sigma)}\, d\sigma \quad \text{for } s,\ x \in \mathbb{R}.$$

Show that, for each $x \in \mathbb{R}$, T_x is a strictly increasing map of $\mathbb{R}$ onto itself, and define $X(t, x) = x + T_x^{-1}(t)$. Also, show that $\dot{X}(t) = F\big(X(t)\big)$ with $X(0) = x$ if and only if $X(t) = X(t, x)$ for all $t \in \mathbb{R}$. In other words, solutions to this equations are simply the path $t \rightsquigarrow x + t$ run with the "clock" $t \rightsquigarrow T_x^{-1}(t)$. The reason why things are so simple here is that, aside from the rate at which it is traveled, there is only one increasing, continuous path in $\mathbb{R}$.

Exercise 4.20 What motivated Fourier to introduce his representation of functions was that he wanted to use it to find solutions to partial differential equations. To see what he had in mind, suppose that $f : [0, 1] \longrightarrow \mathbb{C}$ is a continuous function, and show that the function $u : [0, 1] \times (0, \infty) \longrightarrow \mathbb{C}$ given by

$$u(x, t) = 2 \sum_{m=1}^{\infty} e^{-(m\pi)^2 t} \left(\int_0^1 f(y) \sin(m\pi y)\, dy \right) \sin(m\pi x)$$

solves the *heat equation* $\partial_t u(x,t) = \partial_x^2 u(x,t)$ in $(0,1) \times (0,\infty)$ with boundary conditions $u(0,t) = 0 = u(1,t)$ and initial condition $\lim_{t \searrow 0} u(x,t) = f(x)$ for $x \in (0,1)$.

Chapter 5
Integration in Higher Dimensions

Integration of functions in higher dimensions is much more difficult than it is in one dimension. The basic reason is that in order to integrate a function, one has to know how to measure the volume of sets. In one dimension, most sets can be decomposed into intervals (cf. Exercise 1.21), and we took the length of an interval to be its volume. However, already in $\mathbb{R}^2$ there is a vastly more diverse menagerie of shapes. Thus knowing how to integrate over one shape does not immediately tell you how to integrate over others. A second reason is that, even if one knows how to define the integral of functions in $\mathbb{R}^N$, in higher dimensions there is no comparable *deus ex machina* to replace The Fundamental Theorem of Calculus.

A thoroughly satisfactory theory that addresses the first issue was developed by Lebesgue, but, because it takes too much time to explain, his is not the theory presented here. Instead, we will stay with Riemann's approach.

5.1 Integration Over Rectangles

The simplest analog in $\mathbb{R}^N$ of a closed interval is a closed rectangle R,[1] a set of the form

$$\prod_{j=1}^{N}[a_j, b_j] = [a_1, b_1] \times \cdots \times [a_N, b_N] = \{\mathbf{x} \in \mathbb{R}^N : a_j \le x_j \le b_j \text{ for } 1 \le j \le N\},$$

where $a_j \le b_j$ for each j. Such rectangles have three great virtues. First, if one includes the empty set $\emptyset$ as a rectangle, then the intersection of any two rectangles is again a rectangle. Secondly, there is no question how to assign the volume $|R|$ of a rectangle, it's got to be $\prod_{j=1}^{N}(b_j - a_j)$, the product of the lengths of its sides. Finally, rectangles are easily subdivided into other rectangles. Indeed, every subdivision of

[1] From now on, every rectangle will be assumed to be closed unless it is explicitly stated that it is not.

D.W. Stroock, *A Concise Introduction to Analysis*,
DOI 10.1007/978-3-319-24469-3_5

the intervals making up its sides leads to a subdivision of the rectangle into sub-rectangles. With this in mind, we will now mimic the procedure that we carried out in Sect. 3.1.

Much of what follows relies on the following, at first sight obvious, lemma. In its statement, and elsewhere, two sets are said to be *non-overlapping* if their interiors are disjoint.

Lemma 5.1.1 *If $\mathcal{C}$ is a finite collection of non-overlapping rectangles each of which is contained in the rectangle R, then $|R| \geq \sum_{S\in\mathcal{C}} |S|$. On the other hand, if $\mathcal{C}$ is any finite collection of rectangles whose union contains a rectangle R, then $|R| \leq \sum_{S\in\mathcal{C}} |S|$.*

Proof Since $|S \cap R| \leq |S|$, we may and will assume throughout that $R \supseteq \bigcup_{S\in\mathcal{C}} S$. Also, without loss in generality, we will assume that $\mathrm{int}(R) \neq \emptyset$.

The proof is by induction on N. Thus, suppose that $N = 1$. Given a closed interval I, use a_I and b_I to denote its left and right endpoints. Determine the points $a_R \leq c_0 < \cdots < c_\ell \leq b_R$ so that

$$\{c_k : 0 \leq k \leq \ell\} = \{a_I : I \in \mathcal{C}\} \cup \{b_I : I \in \mathcal{C}\},$$

and set $\mathcal{C}_k = \{I \in \mathcal{C} : [c_{k-1}, c_k] \subseteq I\}$. Clearly $|I| = \sum_{\{k:\, I\in\mathcal{C}_k\}}(c_k - c_{k-1})$ for each $I \in \mathcal{C}$.[2]

When the intervals in $\mathcal{C}$ are non-overlapping, no $\mathcal{C}_k$ contains more than one $I \in \mathcal{C}$, and so

$$\begin{aligned}\sum_{I\in\mathcal{C}} |I| &= \sum_{I\in\mathcal{C}} \sum_{\{k:I\in\mathcal{C}_k\}} (c_k - c_{k-1}) = \sum_{k=1}^{\ell} \mathrm{card}(\mathcal{C}_k)(c_k - c_{k-1}) \\ &\leq \sum_{k=1}^{\ell} (c_k - c_{k-1}) \leq (b_R - a_R) = |R|.\end{aligned}$$

If $R = \bigcup_{I\in\mathcal{C}} I$, then $c_0 = a_R$, $c_\ell = b_R$, and, for each $0 \leq k \leq \ell$, there is an $I \in \mathcal{C}$ for which $I \in \mathcal{C}_k$. To prove this last assertion, simply note that if $x \in (c_{k-1}, c_k)$ and $\mathcal{C} \ni I \ni x$, then $[c_{k-1}, c_k] \subseteq I$ and therefore $I \in \mathcal{C}_k$. Knowing this, we have

$$\begin{aligned}\sum_{I\in\mathcal{C}} |I| &= \sum_{I\in\mathcal{C}} \sum_{\{k:I\in\mathcal{C}_k\}} (c_k - c_{k-1}) = \sum_{k=1}^{\ell} \mathrm{card}(\mathcal{C}_k)(c_k - c_{k-1}) \\ &\geq \sum_{k=1}^{\ell} (c_k - c_{k-1}) = (b_R - a_R) = |R|.\end{aligned}$$

[2] Here, and elsewhere, the sum over the empty set is taken to be 0.

Now assume the result for N. Given a rectangle S in $\mathbb{R}^{N+1}$, determine $a_S,\ b_S \in \mathbb{R}$ and the rectangle Q_S in $\mathbb{R}^N$ so that $S = Q_S \times [a_S, b_S]$. As before, choose points $a_R \le c_0 < \cdots < c_\ell \le b_R$ for $\{[a_S, b_S] : S \in \mathcal{C}\}$, and define

$$\mathcal{C}_k = \{S \in \mathcal{C} : [c_{k-1}, c_k] \subseteq [a_S, b_S]\}.$$

Then, for each $S \in \mathcal{C}$,

$$|S| = |Q_S|(b_S - a_S) = |Q_S| \sum_{\{k: S\in\mathcal{C}_k\}} (c_k - c_{k-1}).$$

If the rectangles in $\mathcal{C}$ are non-overlapping, then, for each k, the rectangles in $\{Q_S : S \in \mathcal{C}_k\}$ are non-overlapping. Hence, since $\bigcup_{S\in\mathcal{C}_k} Q_S \subseteq Q_R$, the induction hypothesis implies $\sum_{S\in\mathcal{C}_k} |Q_S| \le |Q_R|$ for each $1 \le k \le \ell$, and therefore

$$\begin{aligned}\sum_{S\in\mathcal{C}} |S| &= \sum_{S\in\mathcal{C}} |Q_S| \sum_{\{k:\, S\in\mathcal{C}_k\}} (c_k - c_{k-1}) \\ &= \sum_{k=1}^{\ell} (c_k - c_{k-1}) \sum_{S\in\mathcal{C}_k} |Q_S| \le (b_R - a_R)|Q_R| = |R|.\end{aligned}$$

Finally, assume that $R = \bigcup_{S\in\mathcal{C}} S$. In this case, $c_0 = a_R$ and $c_\ell = b_R$. In addition, for each $1 \le k \le \ell$, $Q_R = \bigcup_{S\in\mathcal{C}_k} Q_S$. To see this, note that if $\mathbf{x} = (x_1, \ldots, x_{N+1}) \in R$ and $x_{N+1} \in (c_{k-1}, c_k)$, then $S \ni \mathbf{x} \implies [c_{k-1}, c_k] \subseteq [a_S, b_S]$ and therefore that $S \in \mathcal{C}_k$. Hence, by the induction hypothesis, $|Q_R| \le \sum_{S\in\mathcal{C}_k} \mathrm{vol}(Q_S)$ for each $1 \le k \le \ell$, and therefore

$$\begin{aligned}\sum_{S\in\mathcal{C}} |S| &= \sum_{S\in\mathcal{C}} |Q_S| \sum_{\{k: S\in\mathcal{C}_k\}} (c_k - c_{k-1}) \\ &= \sum_{k=1}^{\ell} (c_k - c_{k-1}) \sum_{S\in\mathcal{C}_k} |Q_S| \ge (b_R - a_R)|Q_R| = |R|.\end{aligned}$$

□

Given a rectangle $\prod_{j=1}^N [a_j, b_j]$, throughout this section $\mathcal{C}$ will be a finite collection of non-overlapping, closed rectangles R whose union is $\prod_{j=1}^N [a_j, b_j]$, and the mesh size $\|\mathcal{C}\|$ will be $\max\{\mathrm{diam}(R) : R \in \mathcal{C}\}$, where the diameter $\mathrm{diam}(R)$ of $R = \prod_{j=1}^N [r_j, s_j]$ equals $\sqrt{\sum_{j=1}^N (s_j - r_j)^2}$. For instance, $\mathcal{C}$ might be obtained by subdividing each of the sides $[a_j, b_j]$ into n equal parts and taking $\mathcal{C}$ to be the set of n^N rectangles

$$\prod_{j=1}^{N}\left[a_j + \tfrac{m_j-1}{n}(b_j - a_j), a_j + \tfrac{m_j}{n}(b_j - a_j)\right] \quad \text{for } 1 \le m_1, \dots, m_N \le n.$$

Next, say that $\Xi : \mathcal{C} \longrightarrow \mathbb{R}^N$ is a *choice function* if $\boldsymbol{\Xi}(R) \in R$ for each $R \in \mathcal{C}$, and define the *Riemann sum*

$$\mathcal{R}(f; \mathcal{C}, \boldsymbol{\Xi}) = \sum_{R \in \mathcal{C}} f\big(\boldsymbol{\Xi}(R)\big)|R|$$

for bounded functions $f : \prod_{j=1}^N [a_j, b_j] \longrightarrow \mathbb{R}$. Again, we say that f is *Riemann integrable* if there exists a $\int_{\prod_1^N [a_j,b_j]} f(\mathbf{x})\, d\mathbf{x} \in \mathbb{R}$ to which the Riemann sums $\mathcal{R}(f; \mathcal{C}, \Xi)$ converge, in the same sense as before, as $\|\mathcal{C}\| \to 0$, in which case $\int_{\prod_1^N [a_j,b_j]} f(\mathbf{x})\, d\mathbf{x}$ is called the *Riemann integral* or just the *integral* of f on $\prod_{j=1}^N [a_j, b_j]$.

There are no essentially new ideas needed to analyze when a function is Riemann integrable. As we did in Sect. 3.1, one introduces the upper and lower Riemann sums

$$\mathcal{U}(f; \mathcal{C}) = \sum_{R \in \mathcal{C}} \left(\sup_R f\right)|R| \quad \text{and} \quad \mathcal{L}(f; \mathcal{C}) = \sum_{R \in \mathcal{C}} \left(\inf_R f\right)|R|,$$

and, using the same reasoning as we did in the proof of Lemma 3.1.1, checks that $\mathcal{L}(f; \mathcal{C}) \le \mathcal{R}(f; \mathcal{C}, \boldsymbol{\Xi}) \le \mathcal{U}(f; \mathcal{C})$ for any $\boldsymbol{\Xi}$ and $\mathcal{L}(f; \mathcal{C}) \le \mathcal{U}(f; \mathcal{C}')$ for any $\mathcal{C}'$. Further, one can show that for each $\mathcal{C}$ and $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\|\mathcal{C}'\| < \delta \implies \mathcal{U}(f; \mathcal{C}') \le \mathcal{U}(f; \mathcal{C}) + \epsilon \text{ and } \mathcal{L}(f; \mathcal{C}') \ge \mathcal{L}(f; \mathcal{C}) - \epsilon.$$

The proof that such a δ exists is basically the same as, but somewhat more involved than, the corresponding one in Lemma 3.1.1. Namely, given $\delta > 0$ and a rectangle $R = \prod_{j=1}^N [c_j, d_j] \in \mathcal{C}$, define $R_k^-(\delta)$ and $R_k^+(\delta)$ to be the rectangles

$$\left(\prod_{1 \le j < k} [a_j, b_j]\right) \times \left[a_k \vee (c_k - \delta), b_k \wedge (c_k + \delta)\right] \times \left(\prod_{k < j \le N} [a_j, b_j]\right)$$

and

$$\left(\prod_{1 \le j < k} [a_j, b_j]\right) \times \left[a_k \vee (d_k - \delta), b_k \wedge (d_k + \delta)\right] \times \left(\prod_{k < j \le N} [a_j, b_j]\right)$$

for $1 \le k \le N$, with the understanding that the first factor is absent if $k = 1$ and the last factor is absent if $k = N$. Now suppose that $\|\mathcal{C}'\| < \delta$ and $R' \in \mathcal{C}'$. Then either $R' \subseteq R$ for some $R \in \mathcal{C}$ or there is an $1 \le k \le N$ and an $R \in \mathcal{C}$ such that the interior of the kth side of R' contains one of the end points of kth side of R, in which case $R' \subseteq R_k^-(\delta) \cup R_k^+(\delta)$. Thus, if $\mathcal{D}$ is the set of $R' \in \mathcal{C}'$ that are not contained in any $R \in \mathcal{C}$, then, because $\sup_{R'} f \le \sup_R f$ if $R' \subseteq R$, one can use Lemma 5.1.1 to see that

$$\begin{aligned}\mathcal{U}(f;\mathcal{C}') - \mathcal{U}(f;\mathcal{C}) &= \sum_{R'\in\mathcal{C}'}\sum_{R\in\mathcal{C}}\left(\sup_{R'} f - \sup_R f\right)|R'\cap R| \\ &\le \sum_{R'\in\mathcal{D}}\sum_{R\in\mathcal{C}}\left(\sup_{R'} f - \sup_R f\right)|R'\cap R| \le 2\|f\|_{\prod_1^N[a_j,b_j]}\sum_{R'\in\mathcal{D}}|R'| \\ &\le 2\|f\|_{\prod_1^N[a_j,b_j]}\sum_{k=1}^N\sum_{R\in\mathcal{C}}\bigl(|R_k^-(\delta)| + |R_k^+(\delta)|\bigr).\end{aligned}$$

Since $|R_k^\pm(\delta)| \le \delta\prod_{j\ne k}(b_j - a_j)$, it follows that there exists a constant $A < \infty$ such that $\mathcal{U}(f;\mathcal{C}') \le \mathcal{U}(f;\mathcal{C}) + A\delta$ if $\|\mathcal{C}'\| < \delta$.

With these preparations, we now have the following analog of Theorem 3.1.2. However, before stating the result, we need to make another definition. Namely, we will say that a subset Γ of the rectangle $\prod_{j=1}^N[a_j,b_j]$ is *Riemann negligible* if, for each $\epsilon > 0$ there is a $\mathcal{C}$ such that

$$\sum_{\substack{R\in\mathcal{C}\\ \Gamma\cap R\ne\emptyset}}|R| < \epsilon.$$

Riemann negligible sets will play an important role in our considerations.

Theorem 5.1.2 *Let $f : \prod_{j=1}^N[a_j,b_j] \longrightarrow \mathbb{C}$ be a bounded function. Then f is Riemann integrable if and only if for each $\epsilon > 0$ there is a $\mathcal{C}$ such that*

$$\sum_{\substack{R\in\mathcal{C}\\ \sup_R f - \inf_R f \ge \epsilon}}|R| < \epsilon.$$

In particular, f is Riemann integrable if it is continuous off of a Riemann negligible set. Finally if f is Riemann integrable and takes all its values in a compact set $K \subseteq \mathbb{C}$ and $\varphi : K \longrightarrow \mathbb{C}$ is continuous, then $\varphi\circ f$ is Riemann integrable.

Proof Except for the one that says f is Riemann integrable if it is continuous off of a Riemann negligible set, all these assertions are proved in exactly the same way as the analogous statements in Theorem 3.1.2.

Now suppose that f is continuous off of the Riemann negligible set Γ. Given $\epsilon > 0$, choose $\mathcal{C}$ so that $\sum_{R\in\mathcal{D}}|R| < \epsilon$, where $\mathcal{D} = \{R \in \mathcal{C} : R\cap\Gamma \ne \emptyset\}$. Then $K = \bigcup_{R\in\mathcal{C}\setminus\mathcal{D}} R$ is a compact set on which f is continuous. Hence, we can find a $\delta > 0$ such that $|f(y) - f(x)| < \epsilon$ for all $x, y \in K$ with $|y - x| \le \delta$. Finally, subdivide each $R \in \mathcal{C}\setminus\mathcal{D}$ into rectangles of diameter less than δ, and take $\mathcal{C}'$ to be the cover consisting of the elements of $\mathcal{D}$ and the sub-rectangles into which the elements of $\mathcal{C}\setminus\mathcal{D}$ were subdivided. Then

$$\sum_{\substack{R'\in\mathcal{C}' \\ \sup_{R'} f - \inf_{R'} f \geq \epsilon}} |R'| \leq \sum_{R\in\mathcal{D}} |R| < \epsilon. \qquad \square$$

We now have the basic facts about Riemann integration in $\mathbb{R}^N$, and from them follow the Riemann integrability of linear combinations and products of bounded Riemann integrable functions as well as the obvious analogs of (3.1.1), (3.1.5), (3.1.4), and Theorem 3.14. The replacement for (3.1.2) is

$$\int_{\prod_1^N [\lambda a_j, \lambda b_j]} f(\mathbf{x})\, d\mathbf{x} = \lambda^N \int_{\prod_1^N [a_j, b_j]} f(\lambda \mathbf{x})\, dx \tag{5.1.1}$$

for bounded, Riemann integrable functions on $\prod_{j=1}^N [\lambda a_j, \lambda b_j]$. It is also useful to note that Riemann integration is *translation invariant* in the sense that if f is a bounded, Riemann integrable function on $\prod_{j=1}^N [c_j + a_j, c_j + b_j]$ for some $\mathbf{c} = (c_1, \dots, c_N) \in \mathbb{R}^N$, then $x \rightsquigarrow f(\mathbf{c} + \mathbf{x})$ is Riemann integrable on $\prod_{j=1}^N [a_j, b_j]$ and

$$\int_{\prod_1^N [c_j + a_k, c_j + b_j]} f(\mathbf{x})\, d\mathbf{x} = \int_{\prod_1^N [a_j, b_j]} f(\mathbf{c} + \mathbf{x})\, d\mathbf{x}, \tag{5.1.2}$$

a property that follows immediately from the corresponding fact for Riemann sums. In addition, by the same procedure as we used in Sect. 3.1, we can extend the definition of the Riemann integral to cover situations in which either the integrand f or the region over which the integration is performed is unbounded. Thus, for example, if f is a function that is bounded and Riemann integrable on bounded rectangles, then one defines

$$\int_{\mathbb{R}^N} f(\mathbf{x})\, d\mathbf{x} = \lim_{\substack{a_1 \vee \cdots \vee a_N \to -\infty \\ b_1 \wedge \cdots \wedge b_N \to \infty}} \int_{\prod_1^N [a_j, b_j]} f(\mathbf{x})\, d\mathbf{x}$$

if the limit exists.

5.2 Iterated Integrals and Fubini's Theorem

Evaluating integrals in N variables is hard and usually possible only if one can reduce the computation to integrals in one variable. One way to make such a reduction is to write an integral in N variables as N iterated integrals in one variable, one for each dimension, and the following theorem, known as *Fubini's Theorem*, shows this can be done. In its statement, if $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$ and $1 \leq M < N$, then $\mathbf{x}_1^{(M)} \equiv (x_1, \dots, x_M)$ and $\mathbf{x}_2^{(M)} \equiv (x_{M+1}, \dots, x_N)$.

Theorem 5.2.1 *Suppose that* $f : \prod_{j=1}^N [a_j, b_j] \longrightarrow \mathbb{C}$ *is a bounded, Riemann integrable function. Further, for some* $1 \le M < N$ *and each* $\mathbf{x}_2^{(M)} \in \prod_{j=M+1}^N [a_j, b_j]$, *assume that* $\mathbf{x}_1^{(M)} \in \prod_{j=1}^M [a_j, b_j] \longmapsto f(\mathbf{x}_1^{(M)}, \mathbf{x}_2^{(M)}) \in \mathbb{C}$ *is Riemann integrable. Then*

$$\mathbf{x}_2^{(M)} \in \prod_{j=M+1}^N [a_j, b_j] \longmapsto f_1^{(M)}(\mathbf{x}_2^{(M)}) \equiv \int_{\prod_1^M [a_j,b_j]} f(\mathbf{x}_1^{(M)}, \mathbf{x}_2^{(M)})\, d\mathbf{x}_1^{(M)}$$

is Riemann integrable and

$$\int_{\prod_1^N [a_j,b_j]} f(\mathbf{x})\, d\mathbf{x} = \int_{\prod_{M+1}^N [a_j,b_j]} f_1^{(M)}(\mathbf{x}_2^{(M)})\, d\mathbf{x}_2^{(M)}.$$

In particular, this result applies if f *is a bounded, Riemann integrable function with the property that, for each* $\mathbf{x}_2^{(M)} \in \prod_{j=M+1}^N [a_j, b_j]$, $\mathbf{x}_1^{(M)} \in \prod_{j=1}^M [a_j, b_j] \longmapsto f(\mathbf{x}_1^{(M)}, \mathbf{x}_2^{(M)}) \in \mathbb{C}$ *is continuous at all but a Riemann negligible set of points.*

Proof Given $\epsilon > 0$, choose $\delta > 0$ so that

$$\|\mathcal{C}\| < \delta \implies \left| \int_{\prod_1^N [a_j,b_j]} f(\mathbf{x})\, dx - \mathcal{R}(f;\, \mathcal{C}, \boldsymbol{\Xi}) \right| < \epsilon$$

for every choice function $\boldsymbol{\Xi}$. Next, let $\mathcal{C}_2^{(M)}$ be a cover of $\prod_{j=M+1}^N [a_j, b_j]$ with $\|\mathcal{C}_2^{(M)}\| < \frac{\delta}{2}$, and let $\boldsymbol{\Xi}_2^{(M)}$ be an associated choice function. Finally, because $\mathbf{x}_1^{(M)} \rightsquigarrow f\big(\mathbf{x}_1^{(M)}, \boldsymbol{\Xi}_2^{(M)}(R_2)\big)$ is Riemann integrable for each $R_2 \in \mathcal{C}_2^{(M)}$, we can choose a cover $\mathcal{C}_1^{(M)}$ of $\prod_{j=1}^M [a_j, b_j]$ with $\|\mathcal{C}_1^{(M)}\| < \frac{\delta}{2}$ and an associated choice function $\boldsymbol{\Xi}_1^{(M)}$ such that

$$\sum_{R_2 \in \mathcal{C}_2^{(M)}} \left| \sum_{R_1 \in \mathcal{C}_1^{(M)}} f\big(\boldsymbol{\Xi}_1^{(M)}(R_1), \boldsymbol{\Xi}_2^{(M)}(R_2)\big)\big)|R_1| - f_1^{(M)}(\boldsymbol{\Xi}_2^{(M)}(R_2)\big) \right| |R_2| < \epsilon.$$

If

$$\mathcal{C} = \{R_1 \times R_2 :\ R_1 \in \mathcal{C}_1^{(M)}\ \&\ R_2 \in \mathcal{C}_2^{(M)}\}$$

and $\boldsymbol{\Xi}\big(R_1 \times R_2\big) = \big(\boldsymbol{\Xi}_1^{(M)}(R_1), \boldsymbol{\Xi}_2^{(M)}(R_2)\big)$, then $\|\mathcal{C}\| < \delta$ and

$$\mathcal{R}(f; \mathcal{C}, \boldsymbol{\Xi}) = \sum_{R_1 \in \mathcal{C}_1^{(M)}} \sum_{R_2 \in \mathcal{C}_2^{(M)}} f\big(\boldsymbol{\Xi}_1^{(M)}(R_1), \boldsymbol{\Xi}_2^{(M)}(R_2)\big)|R_1||R_2|,$$

and so

$$\begin{aligned}&\left|\int_{\prod_1^N[a_j,b_j]} f(\mathbf{x})\,d\mathbf{x} - \mathcal{R}(f_1^{(M)};\mathcal{C}_2^{(M)},\boldsymbol{\Xi}_2^{(M)})\right|\\ &\le \left|\int_{\prod_1^N[a_j,b_j]} f(\mathbf{x})\,d\mathbf{x} - \mathcal{R}(f;\mathcal{C},\boldsymbol{\Xi})\right|\\ &\quad + \sum_{R_2\in\mathcal{C}_2^{(M)}}\left|\sum_{R_1\in\mathcal{C}_1^{(M)}} f\big(\boldsymbol{\Xi}_1^{(M)}(R_1),\boldsymbol{\Xi}_2^{(M)}(R_2)\big)|R_1| - f_1^{(M)}(\boldsymbol{\Xi}_2^{(M)}(R_2))\right||R_2|\end{aligned}$$

is less than 2ϵ. Hence, $\mathcal{R}(f_1^{(M)};\mathcal{C}_2^{(M)},\boldsymbol{\Xi}_2^{(M)}\big)$ converges to $\int_{\prod_1^N[a_j,b_j]} f(\mathbf{x})\,d\mathbf{x}$ as $\|\mathcal{C}_2^{(M)}\|\to 0$. □

It should be clear that the preceding result holds equally well when the roles of $\mathbf{x}_1^{(M)}$ and $\mathbf{x}_2^{(M)}$ are reversed. Thus, if $f:\prod_{j=1}^N[a_j,b_j]\longmapsto\mathbb{C}$ is a bounded, Riemann integrable function such that $\mathbf{x}_1^{(M)}\in\prod_{j=1}^M[a_j,b_j]\longmapsto f(\mathbf{x}_1^{(M)},\mathbf{x}_2^{(M)})\in\mathbb{C}$ is Riemann integrable for each $\mathbf{x}_2^{(M)}\in\prod_{j=M+1}^N[a_j,b_j]$ and $\mathbf{x}_2^{(M)}\in\prod_{j=M+1}^N[a_j,b_j]\longmapsto f(\mathbf{x}_1^{(M)},\mathbf{x}_2^{(M)})$ is Riemann integrable for each $\mathbf{x}_1^{(M)}\in\prod_{j=1}^N[a_j,b_j]$, then

$$\int_{\prod_1^N[a_j,b_j]} f(\mathbf{x})\,d\mathbf{x} = \begin{cases}\int_{\prod_{M+1}^N[a_j,b_j]} f_1^{(M)}(\mathbf{x}_2^{(M)})\,d\mathbf{x}_2^{(M)}\\ \int_{\prod_1^M[a_j,b_j]} f_2^{(M)}(\mathbf{x}_2^{(M)})\,d\mathbf{x}_1^{(M)},\end{cases} \tag{5.2.1}$$

where

$$f_1^{(M)}(\mathbf{x}_2^{(M)}) = \int_{\prod_1^M[a_j,b_j]} f(\mathbf{x}_1^{(M)},\mathbf{x}_2^{(M)})\,d\mathbf{x}_1^{(M)}$$

and

$$f_2^{(M)}(\mathbf{x}_1^{(M)}) = \int_{\prod_{M+1}^N[a_j,b_j]} f(\mathbf{x}_1^{(M)},\mathbf{x}_2^{(M)})\,d\mathbf{x}_2^{(M)}.$$

Corollary 5.2.2 *Let f be a continuous function on $\prod_{j=1}^N[a_j,b_j]$. Then for each $1\le M<N$,*

$$\mathbf{x}_2^{(M)}\in\prod_{j=M+1}^N[a_j,b_j]\longmapsto f_1^{(M)}(\mathbf{x}_2^{(M)})\equiv\int_{\prod_1^M[a_j,b_j]} f(\mathbf{x}_1^{(M)},\mathbf{x}_2^{(M)})\,d\mathbf{x}_1^{(M)}\in\mathbb{C}$$

is continuous. Furthermore,

$$f_1^{(M+1)}(\mathbf{x}_2^{(M+1)}) = \int_{[a_M,b_M]} f_1^{(M)}(x_M,\mathbf{x}_2^{M+1})\,dx_M \quad \textit{for } 1\le M<N-1$$

and

$$\int_{\prod_1^N [a_j,b_j]} f(\mathbf{x})\,d\mathbf{x} = \int_{[a_N,b_N]} f_1^{(N-1)}(x_N)\,dx_N.$$

Proof Once the first assertion is proved, the others follow immediately from Theorem 5.2.1. But, because f is uniformly continuous, the first assertion follows from the obvious higher dimensional analog of Theorem 3.1.4. □

By repeated applications of Corollary 5.2.2, one sees that

$$\int_{\prod_{j=1}^N [a_j,b_j]} f(\mathbf{x})\,d\mathbf{x} = \int_{a_N}^{b_N} \left(\cdots \left(\int_{a_1}^{b_1} f(x_1,\dots,x_{N-1},x_N)\,dx_1 \right) \cdots \right) dx_N.$$

The expression on the right is called an *iterated integral*. Of course, there is nothing sacrosanct about the order in which one does the integrals. Thus

$$\int_{\prod_{j=1}^N [a_j,b_j]} f(\mathbf{x})\,d\mathbf{x} = \int_{a_{\pi(N)}}^{b_{\pi(N)}} \left(\cdots \left(\int_{a_{\pi(1)}}^{b_{\pi(1)}} f(x_1,\dots,x_{N-1},x_N)\,dx_{\pi(1)} \right) \cdots \right) dx_{\pi(N)} \tag{5.2.2}$$

for any permutation π of $\{1,\dots,N\}$. In that it shows integrals in N variables can be evaluated by doing N integrals in one variable, (5.2.2) makes it possible to bring Theorem 3.2.1 to bear on the problem. However, it is hard enough to find one indefinite integral on $\mathbb{R}$, much less a succession of N of them. Nonetheless, there is an important consequence of (5.2.2). Namely, if $f(\mathbf{x}) = \prod_{j=1}^N f_j(x_j)$, where, for each $1 \le j \le N$, f_j is a continuous function on $[a_j,b_j]$, then

$$\int_{\prod_1^N [a_j,b_j]} f(\mathbf{x})\,d\mathbf{x} = \prod_{j=1}^N \int_{[a_j,b_j]} f_j(x_j)\,dx_j. \tag{5.2.3}$$

In fact, starting from Theorem 5.2.1, it is easy to check that (5.2.3) holds when each f_j is bounded and Riemann integrable.

Looking at (5.2.2), one might be tempted to think that there is an analog of the Fundamental Theorem of Calculus for integrals in several variables. Namely, taking π to be the identity permutation that leaves the order unchanged and thinking of the expression on the right as a function F of $(b_1,\dots,b_N)$, it becomes clear that $\partial_{\mathbf{e}_1}\dots\partial_{\mathbf{e}_N}F = f$. However, what made this information valuable when $N=1$ is the fact that a function on $\mathbb{R}$ can be recovered, up to an additive constant, from its derivative, and that is why we could say that $F(b) - F(a) = \int_a^b f(x)\,dx$ for any

F satisfying $F' = f$. When $N \geq 2$, the equality $\partial_{\mathbf{e}_1} \ldots \partial_{\mathbf{e}_N} F = f$ provides much less information. Indeed, even when $N = 2$, if F satisfies $\partial_{\mathbf{e}_1}\partial_{\mathbf{e}_2} F = f$, then so does $F(x_1, x_2) + F_1(x_1) + F_2(x_2)$ for any choice of differentiable functions F_1 and F_2, and the ambiguity gets worse as N increases. Thus finding an F that satisfies $\partial_{\mathbf{e}_1} \ldots \partial_{\mathbf{e}_N} F = f$ does little to advance one toward finding the integral of f.

To provide an interesting example of the way in which Fubini's Theorem plays an important role, define *Euler's Beta function* $B : (0, \infty)^2 \longrightarrow (0, \infty)$ by

$$B(\alpha, \beta) = \int_{(0,1)} x^{\alpha-1}(1-x)^{\beta-1}\, dx.$$

It turns out that his Beta function is intimately related to his (cf. Exercise 3.3) Gamma function. In fact,

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \tag{5.2.4}$$

which means that $\frac{1}{B(\alpha,\beta)}$ is closely related to the binomial coefficients in the same sense that $\Gamma(t)$ is related to factorials. Although (5.2.4) holds for all $(\alpha, \beta) \in (0, \infty)^2$, in order to avoid distracting technicalities, we will prove it only for $(\alpha, \beta) \in [1, \infty)^2$. Thus let $\alpha, \beta \geq 1$ be given. Then, by (5.2.3) and (5.2.1), $\Gamma(\alpha)\Gamma(\beta)$ equals

$$\begin{aligned}&\lim_{r\to\infty} \int_{[0,r]^2} x_1^{\alpha-1} x_2^{\beta-1} e^{-x_1-x_2}\, d\mathbf{x}\\ &\qquad = \lim_{r\to\infty} \int_0^r x_2^{\beta-1} \left(\int_0^r x_1^{\alpha-1} e^{-(x_1+x_2)}\, dx_1 \right) dx_2.\end{aligned}$$

By (5.1.2),

$$\int_0^r x_1^{\alpha-1} e^{-(x_1+x_2)}\, dx_1 = \int_{x_2}^{r+x_2} (y_1 - x_2)^{\alpha-1} e^{-y_1}\, dy_1,$$

and so

$$\begin{aligned}&\int_0^r x_2^{\beta-1} \left(\int_0^r x_1^{\alpha-1} e^{-(x_1+x_2)}\, dx_1 \right) dx_2\\ &\qquad = \int_0^r x_2^{\beta-1} \left(\int_{x_2}^{r+x_2} (y_1 - x_2)^{\alpha-1} e^{-y_1}\, dy_1 \right) dx_2.\end{aligned}$$

Now consider the function

$$f(y_1, x_2) = \begin{cases} (y_1 - x_2)^{\alpha-1} x_2^{\beta-1} e^{-y_1} & \text{if } x_2 \in [0, r]\ \&\ x_2 \leq y_1 \leq r + x_2 \\ 0 & \text{otherwise} \end{cases}$$

on $[0, 2r]\times[0, r]$. Because the only discontinuities of f lie in the Riemann negligible set $\{(r+x_2, x_2) : x_2 \in [0, r]\}$, it is Riemann integrable on $[0, 2r]\times[0, r]$. In addition, for each $y_1 \in [0, 2r]$, $x_2 \rightsquigarrow f(y_1, x_2)$ and, for each $x_2 \in [0, r]$, $y_1 \rightsquigarrow f(y_1, x_2)$ have at most two discontinuities. We can therefore apply (5.2.1) to justify

$$\int_0^r x_2^{\beta-1}\left(\int_{x_2}^{r+x_2}(y_1-x_2)^{\alpha-1}e^{-y_1}\,dy_1\right)dx_2 = \int_0^r\left(\int_0^{2r} f(y_1,x_2)\,dy_1\right)dx_2$$
$$= \int_0^{2r}\left(\int_0^r f(y_1,x_2)\,dx_2\right)dy_1 = \int_0^{2r} e^{-y_1}\left(\int\limits_{(y_1-r)^+}^{r\wedge y_1}(y_1-x_2)^{\alpha-1}x_2^{\beta-1}\,dx_2\right)dy_1.$$

Further, by (3.1.2)

$$\int_{(y_1-r)^+}^{r\wedge y_1}(y_1-x_2)^{\alpha-1}x_2^{\beta-1}\,dx_2 = y_1^{\alpha+\beta-1}\int_{(1-y_1^{-1}r)^+}^{1\wedge(y_1^{-1}r)}(1-y_2)^{\alpha-1}y_2^{\beta-1}\,dy_2.$$

Collecting these together, we have

$$\Gamma(\alpha)\Gamma(\beta) = \lim_{r\to\infty}\int_0^{2r} y_1^{\alpha+\beta-1}e^{-y_1}\left(\int_{(1-y_1^{-1}r)^+}^{1\wedge(y_1^{-1}r)}(1-y_2)^{\alpha-1}y_2^{\beta-1}\,dy_2\right)dy_1.$$

Finally,

$$\int_0^{2r} y_1^{\alpha+\beta-1}e^{-y_1}\left(\int_{(1-y_1^{-1}r)^+}^{1\wedge(y_1^{-1}r)}(1-y_2)^{\alpha-1}y_2^{\beta-1}\,dy_2\right)dy_1$$
$$= \int_0^r y_1^{\alpha+\beta-1}e^{-y_1}\,dy_1 B(\alpha,\beta)$$
$$+ \int_r^{2r} y_1^{\alpha+\beta-1}e^{-y_1}\left(\int_{(1-y_1^{-1}r)}^{y_1^{-1}r}(1-y_2)^{\alpha-1}y_2^{\beta-1}\,dy_2\right)dy_1,$$

and, as $r \to \infty$, the first term on the right tends to $\Gamma(\alpha+\beta)$ whereas the second term is dominated by $\int_r^\infty y_1^{\alpha+\beta-1}e^{-y_1}\,dy_1$ and therefore tends to 0.

The preceding computation illustrates one of the trickier aspects of proper applications of Fubini's Theorem. When one reverses the order of integration, it is very important to figure out what are the resulting correct limits of integration. As in the application above, the correct limits can look very different after the order of integration is changed.

The Eq. (5.2.4) provides a proof of *Stirling's formula* for the Gamma function as a consequence of (1.8.7). Indeed, by (5.2.4), $\Gamma(n+1+\theta) = \frac{n!\Gamma(\theta)}{B(n+1,\theta)}$ for $n \in \mathbb{Z}^+$ and $\theta \in [1, 2)$, and, by (3.1.2),

$$B(n+1,\theta) = n^{-\theta}\int_0^n y^{\theta-1}\big(1-\tfrac{y}{n}\big)^n\,dy.$$

Further, because $1-x \le e^{-x}$ for all $x \in \mathbb{R}$,

$$\int_0^n y^{\theta-1}\big(1-\tfrac{y}{n}\big)^n\,dy \le \int_0^\infty y^{\theta-1}e^{-y}\,dy = \Gamma(\theta),$$

and, for all $r > 0$,

$$\varliminf_{n\to\infty}\int_0^n y^{\theta-1}\big(1-\tfrac{y}{n}\big)^n\,dy \ge \varliminf_{n\to\infty}\int_0^r y^{\theta-1}\big(1-\tfrac{y}{n}\big)^n\,dy = \int_0^r y^{\theta-1}e^{-y}\,dy.$$

Since the final expression tends to $\Gamma(\theta)$ as $r \to \infty$ uniformly fast for $\theta \in [1,2]$, we now know that

$$\frac{\Gamma(\theta)}{n^\theta B(n+1,\theta)} \longrightarrow 1$$

uniformly fast for $\theta \in [1,2]$. Combining this with (1.8.7) we see that

$$\lim_{n\to\infty}\frac{\Gamma(n+\theta+1)}{\sqrt{2\pi n}\,\big(\frac{n}{e}\big)^n n^\theta} \longrightarrow 1$$

uniformly fast for $\theta \in [1,2]$. Given $t \ge 3$, determine $n_t \in \mathbb{Z}^+$ and $\theta_t \in [1,2)$ so that $t = n_t + \theta_t$. Then the preceding says that

$$\lim_{t\to\infty}\frac{\Gamma(t+1)}{\sqrt{2\pi t}\,\big(\frac{t}{e}\big)^t}\sqrt{\frac{t}{t-\theta_t}}\left(\frac{t}{t-\theta_t}\right)^t e^{-\theta_t} = 1.$$

Finally, it is obvious that, as $t \to \infty$, $\sqrt{\frac{t}{t-\theta_t}}$ tends to 1 and, because, by (1.7.5),

$$\log\left(\left(\frac{t}{t-\theta_t}\right)^t e^{-\theta_t}\right) = -t\log\left(1-\tfrac{\theta_t}{t}\right) - \theta_t \longrightarrow 0,$$

so does $\left(\frac{t}{t-\theta_t}\right)^t e^{-\theta_t}$. Hence we have shown that

$$\Gamma(t+1) \sim \sqrt{2\pi t}\left(\frac{t}{e}\right)^t \quad \text{as } t \to \infty \tag{5.2.5}$$

in the sense that $\lim_{t\to\infty}\frac{\Gamma(t+1)}{\sqrt{2\pi t}\left(\frac{t}{e}\right)^t} = 1$.

5.3 Volume of and Integration Over Sets

We motivated our initial discussion of integration by computing the area under the graph of a non-negative function, and as we will see in this section, integration provides a method for computing the volume of more general regions. However, before we begin, we must first be more precise about what we will mean by the volume of a region.

Although we do not know yet what the volume of a general set Γ is, we know a few properties that volume should possess. In particular, we know that the volume of a subset should be no larger than that of the set containing it. In addition, volume should be additive in the sense that the volume of the union of disjoint sets should be the sum of their volumes. Taking these comments into account, for a given bounded set $\Gamma \subseteq \mathbb{R}^N$, we define the *exterior volume* $|\Gamma|_e$ of Γ to be the infimum of the sums $\sum_{R\in\mathcal{C}} |R|$ as $\mathcal{C}$ runs over all finite collections of non-overlapping rectangles whose union contains Γ.[3] Similarly, define the *interior volume* $|\Gamma|_i$ to be the supremum of the sums $\sum_{R\in\mathcal{C}} |R|$ as $\mathcal{C}$ runs over finite collections of non-overlapping rectangles each of which is contained in Γ. Clearly the notion of exterior volume is consistent with the properties that we want volume to have. To see that the same is true of interior volume, note that an equivalent description would have been that $|\Gamma|_i$ is the supremum of $\sum_{R\in\mathcal{C}} |R|$ as $\mathcal{C}$ runs over finite collections of rectangles that are mutually disjoint and each of which is contained in Γ. Indeed, given a $\mathcal{C}$ of the sort in the definition of interior volume, shrink the sides of each $R \in \mathcal{C}$ with $|R| > 0$ by a factor $\theta \in (0, 1)$ and eliminate the ones with $|R| = 0$. The resulting rectangles will be mutually disjoint and the sum of their volumes will be θ^N times that of the original ones. Hence, by taking θ close enough to 1, we can get arbitrarily close to the original sum.

Obviously Γ is Riemann negligible if and only if $|\Gamma|_e = 0$. Next notice that $|\Gamma|_i \le |\Gamma|_e$ for all bounded Γ's. Indeed, suppose that $\mathcal{C}_1$ is a finite collection of non-overlapping rectangles contained in Γ and that $\mathcal{C}_2$ is a finite collection of rectangles whose union contains Γ. Then, by Lemma 5.1.1,

$$\sum_{R_2\in\mathcal{C}_2} |R_2| \ge \sum_{R_2\in\mathcal{C}_2} \sum_{R_1\in\mathcal{C}_1} |R_1 \cap R_2| = \sum_{R_1\in\mathcal{C}_1} \sum_{R_2\in\mathcal{C}_2} |R_1 \cap R_2| \ge \sum_{R_1\in\mathcal{C}_1} |R_1|.$$

In addition, it is easy to check that

$$\begin{aligned} |\Gamma_1|_e \le |\Gamma_2|_e \text{ and } |\Gamma_1|_i \le |\Gamma_2|_i \quad &\text{if } \Gamma_1 \subseteq \Gamma_2 \\ |\Gamma_1 \cup \Gamma_2|_e \le |\Gamma_1|_e + |\Gamma_2|_e \text{ for all } &\Gamma_1 \;\&\; \Gamma_2, \\ \text{and } |\Gamma_1 \cup \Gamma_2|_i \ge |\Gamma_1|_i + |\Gamma_2|_i \text{ if } &\Gamma_1 \cap \Gamma_2 = \emptyset. \end{aligned}$$

[3] It is reasonable easy to show that $|\Gamma|_e$ would be the same if the infimum were taken over covers by rectangles that are not necessarily non-overlapping.

We will say that Γ is *Riemann measurable* if $|\Gamma|_{\mathrm{i}} = |\Gamma|_{\mathrm{e}}$, in which case we will call $\mathrm{vol}(\Gamma) \equiv |\Gamma|_{\mathrm{e}}$ the *volume* of Γ. Clearly $|\Gamma|_{\mathrm{e}} = 0$ implies that Γ is Riemann measurable and $\mathrm{vol}(\Gamma) = 0$. In particular, if Γ is Riemann negligible and therefore $|\Gamma|_{\mathrm{e}} = 0$, then Γ is Riemann measurable and has volume 0. In addition, if R is a rectangle, $|R|_{\mathrm{e}} \le |R| \le |R|_{\mathrm{i}}$, and therefore R is Riemann measurable and $\mathrm{vol}(R) = |R|$.

One suspects that these considerations are intimately related to Riemann integration, and the following theorem justifies that suspicion. In its statement and elsewhere, $\mathbf{1}_\Gamma$ denotes the *indicator function* of a set Γ. That is, $\mathbf{1}_\Gamma(\mathbf{x})$ is 1 if $\mathbf{x} \in \Gamma$ and is 0 if $\mathbf{x} \notin \Gamma$.

Theorem 5.3.1 *Let Γ be a subset of $\prod_{j=1}^N [a_j, b_j]$. Then Γ is Riemann measurable if and only if $\mathbf{1}_\Gamma$ is Riemann integrable on $\prod_{j=1}^N [a_j, b_j]$, in which case*

$$\mathrm{vol}(\Gamma) = \int_{\prod_{j=1}^N [a_j,b_j]} \mathbf{1}_\Gamma(\mathbf{x})\, d\mathbf{x}.$$

Proof First observe that, without loss in generality, we may assume that all the collections $\mathcal{C}$ entering the definitions of outer and inner volume can be taken to be subsets of non-overlapping covers of $\prod_{j=1}^N [a_j, b_j]$.

Now suppose that Γ is Riemann measurable. Given $\epsilon > 0$, choose a non-overlapping cover $\mathcal{C}_1$ of $\prod_{j=1}^N [a_j, b_j]$ such that

$$\sum_{\substack{R \in \mathcal{C}_1 \\ R \cap \Gamma \neq \emptyset}} |R| \le \mathrm{vol}(\Gamma) + \tfrac{\epsilon}{2}.$$

Then

$$\mathcal{U}(\mathbf{1}_\Gamma; \mathcal{C}_1) = \sum_{\substack{R \in \mathcal{C}_1 \\ R \cap \Gamma \neq \emptyset}} |R| \le \mathrm{vol}(\Gamma) + \tfrac{\epsilon}{2}.$$

Next, choose $\mathcal{C}_2$ so that

$$\sum_{\substack{R \in \mathcal{C}_2 \\ R \subseteq \Gamma}} |R| \ge \mathrm{vol}(\Gamma) - \tfrac{\epsilon}{2},$$

and observe that then $\mathcal{L}(\mathbf{1}_\Gamma; \mathcal{C}_2) \ge \mathrm{vol}(\Gamma) - \frac{\epsilon}{2}$. Hence if

$$\mathcal{C} = \{R_1 \cap R_2 : \ R_1 \in \mathcal{C}_1 \ \&\ R_2 \in \mathcal{C}_2\},$$

then

$$\mathcal{U}(\mathbf{1}_\Gamma; \mathcal{C}) \le \mathcal{U}(\mathbf{1}_\Gamma; \mathcal{C}_1) \le \mathrm{vol}(\Gamma) + \tfrac{\epsilon}{2} \le \mathcal{L}(\mathbf{1}_\Gamma; \mathcal{C}_2) + \epsilon \le \mathcal{L}(\mathbf{1}_\Gamma; \mathcal{C}) + \epsilon,$$

and so not only is $\mathbf{1}_\Gamma$ Riemann integrable but also its integral is equal to $\mathrm{vol}(\Gamma)$.

Conversely, if $\mathbf{1}_\Gamma$ is Riemann integrable and $\epsilon > 0$, choose $\mathcal{C}$ so that $\mathcal{U}(\mathbf{1}_\Gamma; \mathcal{C}) \le \mathcal{L}(\mathbf{1}_\Gamma; \mathcal{C}) + \epsilon$. Define associated choice functions $\boldsymbol{\Xi}_1$ and $\boldsymbol{\Xi}_2$ so that $\boldsymbol{\Xi}_1(R) \in \Gamma$ if $R \cap \Gamma \neq \emptyset$ and $\boldsymbol{\Xi}_2(R) \notin \Gamma$ unless $R \subseteq \Gamma$. Then

$$|\Gamma|_{\mathrm{e}} \le \sum_{\substack{R\in\mathcal{C}\\ R\cap\Gamma\neq\emptyset}} |R| = \mathfrak{R}(\mathbf{1}_\Gamma; \mathcal{C}, \boldsymbol{\Xi}_1) \le \mathfrak{R}(\mathbf{1}_\Gamma; \mathcal{C}, \boldsymbol{\Xi}_2) + \epsilon = \sum_{\substack{R\in\mathcal{C}\\ R\subseteq\Gamma}} |R| + \epsilon \le |\Gamma|_{\mathrm{i}} + \epsilon,$$

and so Γ is Riemann measurable. □

Corollary 5.3.2 *If Γ_1 and Γ_2 are bounded, Riemann measurable sets, then so are $\Gamma_1 \cup \Gamma_2$, $\Gamma_1 \cap \Gamma_2$, and $\Gamma_2 \setminus \Gamma_1$. In addition,*

$$\mathrm{vol}\big(\Gamma_1 \cup \Gamma_2) = \mathrm{vol}(\Gamma_1) + \mathrm{vol}(\Gamma_2) - \mathrm{vol}(\Gamma_1 \cap \Gamma_2)$$

and

$$\mathrm{vol}\big(\Gamma_2 \setminus \Gamma_1\big) = \mathrm{vol}\big(\Gamma_2\big) - \mathrm{vol}\big(\Gamma_1 \cap \Gamma_2\big).$$

In particular, if $\mathrm{vol}(\Gamma_1 \cap \Gamma_2) = 0$, *then* $\mathrm{vol}(\Gamma_1 \cup \Gamma_2) = \mathrm{vol}(\Gamma_1) + \mathrm{vol}(\Gamma_2)$. *Finally, $\Gamma \subseteq \prod_{j=1}^N [a_j, b_j]$ is Riemann measurable if and only if for each $\epsilon > 0$ there exist Riemann measurable subsets A and B of $\prod_{j=1}^N [a_j, b_j]$ such that $A \subseteq \Gamma \subseteq B$ and* $\mathrm{vol}(B \setminus A) < \epsilon$.

Proof By Theorem 5.3.1, $\mathbf{1}_{\Gamma_1}$ and $\mathbf{1}_{\Gamma_2}$ are Riemann integrable. Thus, since

$$\mathbf{1}_{\Gamma_1 \cap \Gamma_2} = \mathbf{1}_{\Gamma_1}\mathbf{1}_{\Gamma_2} \text{ and } \mathbf{1}_{\Gamma_1 \cup \Gamma_2} = \mathbf{1}_{\Gamma_1} + \mathbf{1}_{\Gamma_2} - \mathbf{1}_{\Gamma_1 \cap \Gamma_2},$$

that same theorem implies that $\Gamma_1 \cup \Gamma_2$ and $\Gamma_1 \cap \Gamma_2$ are Riemann measurable. At the same time,

$$\mathbf{1}_{\Gamma_2 \setminus \Gamma_1} = \mathbf{1}_{\Gamma_2} - \mathbf{1}_{\Gamma_1 \cap \Gamma_2},$$

and so $\Gamma_2 \setminus \Gamma_1$ is also Riemann measurable. Also, by Theorem 5.3.1, the equations relating their volumes follow immediately for the equations relating their indicator functions.

Turning to the final assertion, there is nothing to do if Γ is Riemann measurable since we can then take $A = \Gamma = B$ for all $\epsilon > 0$. Now suppose that for each $\epsilon > 0$ there exist Riemann measurable sets A_ϵ and B_ϵ such that $A_\epsilon \subseteq \Gamma \subseteq B_\epsilon \subseteq \prod_{j=1}^N [a_j, b_j]$ and $\mathrm{vol}(B_\epsilon \setminus A_\epsilon) < \epsilon$. Then

$$|\Gamma|_{\mathrm{i}} \ge \mathrm{vol}(A_\epsilon) \ge \mathrm{vol}(B_\epsilon) - \epsilon \ge |\Gamma|_{\mathrm{e}} - \epsilon,$$

and so Γ is Riemann measurable. □

It is reassuring that the preceding result is consistent with our earlier computation of the area under a graph. In fact, we now have the following more general result.

Theorem 5.3.3 *Assume that* $f : \prod_{j=1}^N [a_j, b_j] \longrightarrow \mathbb{R}$ *is continuous. Then the graph*

$$G(f) = \left\{ \big(\mathbf{x}, f(\mathbf{x})\big) : \mathbf{x} \in \prod_{j=1}^N [a_j, b_j] \right\}$$

is a Riemann negligible subset of $\mathbb{R}^{N+1}$. *Moreover, if, in addition,* f *is non-negative and* $\Gamma = \big\{(\mathbf{x}, y) \in \mathbb{R}^{N+1} : 0 \le y \le f(\mathbf{x})\big\}$, *then* Γ *is Riemann measurable and*

$$\mathrm{vol}(\Gamma) = \int_{\prod_1^N [a_j,b_j]} f(\mathbf{x})\, d\mathbf{x}.$$

Proof Set $r = \|f\|_{\prod_1^N [a_j,b_j]}$, and for each $\epsilon > 0$ choose $\delta_\epsilon > 0$ so that

$$|f(\mathbf{y}) - f(\mathbf{x})| < \epsilon \quad \text{if } |\mathbf{y} - \mathbf{x}| \le \delta_\epsilon.$$

Next let $\mathcal{C}$ with $\|\mathcal{C}\| < \delta_\epsilon$ be a cover of $\prod_1^N [a_j, b_j]$ by non-overlapping rectangles, and choose $K \in \mathbb{Z}^+$ so that $\frac{r}{K+1} < \epsilon \le \frac{r}{K}$. Then for each $R \in \mathcal{C}$ there is a $1 \le k_R \le 2(K-1)$ such that

$$\big\{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in R\big\} \subseteq R \times \big[-r + \tfrac{(k_R-1)r}{K}, -r + \tfrac{(k_R+2)r}{K}\big],$$

and therefore

$$|G(f)|_{\mathrm{e}} \le \frac{3r}{K} \sum_{R \in \mathcal{C}} |R| \le 6 \left(\prod_{j=1}^N (b_j - a_j) \right) \epsilon,$$

which proves that $G(f)$ is Riemann negligible.

Turning to the second assertion, note that all the discontinuities of $\mathbf{1}_\Gamma$ on $\prod_{j=1}^N [a_j, b_j] \times [0, r]$ are contained in $G(f)$, and therefore $\mathbf{1}_\Gamma$ is Riemann measurable. In addition, for each $\mathbf{x} \in \prod_{j=1}^N [a_j, b_j]$, $y \in [0, r] \longmapsto \mathbf{1}_\Gamma(\mathbf{x}, y) \in \{0, 1\}$ has at most one discontinuity. Hence, by Theorem 5.3.1 and (5.2.1),

$$\mathrm{vol}(\Gamma) = \int_{\prod_1^N [a_j,b_j]} \left(\int_0^r \mathbf{1}_\Gamma(\mathbf{x}, y)\, dy \right) d\mathbf{x} = \int_{\prod_1^N [a_j,b_j]} f(\mathbf{x})\, d\mathbf{x}. \qquad \square$$

Theorem 5.3.3 allows us to confirm that the volume (i.e., the area) of the closed unit ball $\overline{B(\mathbf{0}, 1)}$ in $\mathbb{R}^2$ is π, the half period of the trigonometric sine function. Indeed, $\overline{B(\mathbf{0}, 1)} = H_+ \cup H_-$, where

$$H_\pm = \left\{ (x_1, x_2) : 0 \le \pm x_2 \le \sqrt{1 - x_1^2} \right\}.$$

By Theorem 5.3.3 both H_+ and H_- are Riemann measurable, and each has area

$$\int_{-1}^{1}\sqrt{(1-x^2)}\,dx = 2\int_0^{\frac{\pi}{2}}\cos^2\theta\,d\theta = \int_0^{\frac{\pi}{2}}\big(1-\cos 2\theta\big)\,d\theta = \frac{\pi}{2}.$$

Finally, $H_+ \cap H_- = [-1,1] \times \{0\}$ is a rectangle with area 0. Hence, by Corollary 5.3.2, the desired conclusion follows. Moreover, because $\mathbf{1}_{\overline{B(\mathbf{0},r)}}(\mathbf{x}) = \mathbf{1}_{\overline{B(\mathbf{0},1)}}(r^{-1}\mathbf{x})$, we can use (5.1.1) and (5.1.2) to see that

$$\mathrm{vol}\big(\overline{B(\mathbf{c},r)}\big) = \pi r^2 \tag{5.3.1}$$

for balls in $\mathbb{R}^2$.

Having defined what we mean by the volume of a set, we now define what we will mean by the integral of a function on a set. Given a bounded, Riemann measurable set Γ, we say that a bounded function $f : \Gamma \longrightarrow \mathbb{C}$ is *Riemann integrable on* Γ if the function

$$\mathbf{1}_\Gamma f \equiv \begin{cases} f & \text{on } \Gamma \\ 0 & \text{off } \Gamma \end{cases}$$

is Riemann integrable on some rectangle $\prod_{j=1}^N [a_j, b_j] \supseteq \Gamma$, in which case the *Riemann integral* of f on Γ is

$$\int_\Gamma f(\mathbf{x})\,d\mathbf{x} \equiv \int_{\prod_1^N [a_j,b_j]} \mathbf{1}_\Gamma(\mathbf{x}) f(\mathbf{x})\,d\mathbf{x}.$$

In particular, if Γ is a bounded, Riemann measurable set, then every bounded, Riemann integrable function on $\prod_{j=1}^N [a_j, b_j]$ will be is Riemann integrable on Γ. In particular, notice that if $\partial\Gamma$ is Riemann negligible and f is a bounded function of Γ that is continuous off of a Riemann negligible set, then f is Riemann integrable on Γ. Obviously, the choice of the rectangle $\prod_{j=1}^N [a_j, b_j]$ is irrelevant as long as it contains Γ.

The following simple result gives an integral version of the intermediate value theorem, Theorem 1.3.6.

Theorem 5.3.4 *Suppose that $K \subseteq \prod_{j=1}^N [a_j, b_j]$ is a compact, connected, Riemann measurable set. If $f : K \longrightarrow \mathbb{R}$ is continuous, then there exists a $\boldsymbol{\xi} \in K$ such that*

$$\int_K f(\mathbf{x})\,dx = f(\boldsymbol{\xi})\mathrm{vol}(K).$$

Proof If $\mathrm{vol}(K) = 0$ there is nothing to do. Now assume that $\mathrm{vol}(K) > 0$. Then $\frac{1}{\mathrm{vol}(K)}\int_K f(\mathbf{x})\,d\mathbf{x}$ lies between the minimum and maximum values that f takes on K, and therefore, by Exercise 4.5 and Lemma 4.1.2, there exists a $\boldsymbol{\xi} \in K$ such that $f(\boldsymbol{\xi}) = \frac{1}{\mathrm{vol}(K)}\int_K f(\mathbf{x})\,d\mathbf{x}$. □

5.4 Integration of Rotationally Invariant Functions

One of the reasons for our introducing the concepts in the preceding section is that they encourage us to get away from rectangles when computing integrals. Indeed, if $\Gamma = \bigcup_{m=0}^{n} \Gamma_m$ where the Γ_m's are bounded Riemann measurable sets, then, for any bounded, Riemann integrable function f,

$$\int_\Gamma f(\mathbf{x})\,d\mathbf{x} = \sum_{m=0}^{n} \int_{\Gamma_m} f(\mathbf{x})\,d\mathbf{x} \quad \text{if } \operatorname{vol}\big(\Gamma_m \cap \Gamma_{m'}\big) = 0 \text{ for } m' \neq m, \tag{5.4.1}$$

since

$$0 \leq \sum_{m=0}^{n} \mathbf{1}_{\Gamma_m} f - \mathbf{1}_\Gamma f \leq 2\|f\|_{\mathrm{u}} \sum_{0\leq m<m'\leq n} \mathbf{1}_{\Gamma_m \cap \Gamma_{m'}}.$$

The advantage afforded by (5.4.1) is that a judicious choice of the Γ_m's can simplify computations. For example, suppose that f is a function on the closed ball $\overline{B(\mathbf{0}, r)}$ in $\mathbb{R}^2$, and assume that $f(\mathbf{x}) = \tilde{f}(|\mathbf{x}|)$ for some continuous function $\tilde{f} : [0, r] \longrightarrow \mathbb{C}$. For each $n \geq 1$, set $\Gamma_{0,n} = \{\mathbf{0}\}$ and $\Gamma_{m,n} = \overline{B\big(\mathbf{0}, \frac{mr}{n}\big)} \setminus \overline{B\big(\mathbf{0}, \frac{(m-1)r}{n}\big)}$ if $1 \leq m \leq n$. By Corollary 5.3.2 and the considerations leading up to (5.3.1), we know that the $\Gamma_{m,n}$'s are Riemann measurable, and, obviously, for each $n \geq 1$ they are a cover of $\overline{B(\mathbf{0}, r)}$ by mutually disjoint sets. If we define

$$f_n(\mathbf{x}) = \sum_{m=0}^{n} \tilde{f}\left(\frac{(2m-1)r}{2n}\right) \mathbf{1}_{\Gamma_{m,n}},$$

then f_n is Riemann measurable, $f_n \longrightarrow f$ uniformly, and therefore

$$\int_{\overline{B(\mathbf{0},r)}} f(\mathbf{x})\,d\mathbf{x} = \lim_{n\to\infty} \int_{\overline{B(\mathbf{0},r)}} f_n(\mathbf{x})\,d\mathbf{x} = \lim_{n\to\infty} \sum_{m=1}^{n} \tilde{f}\big(\tfrac{(2m-1)r}{2n}\big) \operatorname{vol}(\Gamma_{m,n}).$$

Finally, by Corollary 5.3.2 and (5.3.1), $\operatorname{vol}(\Gamma_{m,n}) = \frac{(2m-1)\pi r^2}{n^2}$, and so

$$\int_{\Gamma_m} f_n(\mathbf{x})\,d\mathbf{x} = \frac{2\pi r}{n} \sum_{m=1}^{n} \tilde{f}\big(\tfrac{(2m-1)r}{2n}\big) \tfrac{(2m-1)r}{2n} = 2\pi \mathcal{R}(g; \mathcal{C}_n, \Xi_n),$$

where $g(\rho) = \rho \tilde{f}(\rho)$, $\mathcal{C}_n = \big\{\big[\frac{(m-1)r}{n}, \frac{mr}{n}\big] : 1 \leq m \leq n\big\}$ and $\Xi_n\big(\big[\frac{(m-1)r}{n}, \frac{mr}{n}\big]\big) = \frac{(2m-1)r}{n}$. Hence, we have now proved that

$$\int_{\overline{B(\mathbf{0},r)}} f(\mathbf{x})\,d\mathbf{x} = 2\pi \int_0^r \tilde{f}(\rho)\rho\,d\rho \quad \text{if } f(\mathbf{x}) = \tilde{f}(|\mathbf{x}|) \tag{5.4.2}$$

when $\tilde{f} : [0, r] \longrightarrow \mathbb{C}$ is continuous. The preceding is an example of how, by taking advantage of symmetry properties, one can sometimes reduce the computation of an integral in higher dimensions to one in lower dimensions. In this example the symmetry was the rotational invariance of both the region of integration and the integrand.

Here is a beautiful application of (5.4.2) to a famous calculation. It is known that the function $x \rightsquigarrow e^{-\frac{x^2}{2}}$ does not admit an indefinite integral that can be written as a concatenation of polynomials, trigonometric functions, and exponentials. Nonetheless, by combining (5.2.1) with (5.4.2), we will now show that

$$\int_{\mathbb{R}} e^{-\frac{x^2}{2}}\,dx = \sqrt{2\pi}. \tag{5.4.3}$$

Given $r > 0$, use (5.2.3) to write

$$\left(\int_{-r}^{r} e^{-\frac{x^2}{2}}\,dx\right)^2 = \int_{-r}^{r}\left(\int_{-r}^{r} e^{-\frac{x_1^2+x_2^2}{2}}\,dx_1\right)dx_2 = \int_{[-r,r]^2} e^{-\frac{|\mathbf{x}|^2}{2}}\,d\mathbf{x}.$$

Next observe that

$$\int_{\overline{B(\mathbf{0},\sqrt{2}r)}} e^{-\frac{|\mathbf{x}|^2}{2}}\,d\mathbf{x} \geq \int_{[-r,r]^2} e^{-\frac{|\mathbf{x}|^2}{2}}\,d\mathbf{x} \geq \int_{\overline{B(\mathbf{0},r)}} e^{-\frac{|\mathbf{x}|^2}{2}}\,d\mathbf{x},$$

and that, by (5.4.2),

$$\int_{\overline{B(\mathbf{0},R)}} e^{-\frac{|\mathbf{x}|^2}{2}}\,d\mathbf{x} = 2\pi\int_0^R e^{-\frac{\rho^2}{2}}\rho\,d\rho = 2\pi\big(1 - e^{-\frac{R^2}{2}}\big).$$

Thus, after letting $r \to \infty$, we arrive at (5.4.3). Once one has (5.4.3), there are lots of other computations which follow. For example, one can compute (cf. Exercise 3.3) $\Gamma\big(\frac{1}{2}\big) = \int_0^\infty x^{-\frac{1}{2}}e^{-x}\,dx$. To this end, make the change of variables $y = (2x)^{\frac{1}{2}}$ to see that

$$\begin{aligned}\Gamma\big(\tfrac{1}{2}\big) &= \lim_{r\to\infty}\int_{r^{-1}}^{r} x^{-\frac{1}{2}}e^{-x}\,dx = \lim_{r\to\infty} 2^{\frac{1}{2}}\int_{(2r)^{-\frac{1}{2}}}^{(2r)^{\frac{1}{2}}} e^{-\frac{y^2}{2}}\,dy\\ &= 2^{\frac{1}{2}}\int_{[0,\infty)} e^{-\frac{y^2}{2}}\,dy = 2^{-\frac{1}{2}}\int_{\mathbb{R}} e^{-\frac{y^2}{2}}\,dy,\end{aligned}$$

and conclude that

$$\Gamma\big(\tfrac{1}{2}\big) = \sqrt{\pi}. \tag{5.4.4}$$

We will now develop the N-dimensional analog of (5.4.2) for other $N \geq 1$. Obviously, the 1-dimensional analog is simply the statement that

$$\int_{-r}^{r} f(x)\,dx = 2\int_0^r f(\rho)\,d\rho$$

for even functions f on $[-r, r]$. Thus, assume that $N \geq 3$, and begin by noting that the closed ball $\overline{B(\mathbf{0}, r)}$ of radius $r \geq 0$ centered at the origin is Riemann measurable. Indeed, $\overline{B(\mathbf{0}, r)}$ is the union of the hemispheres

$$H_+ \equiv \left\{\mathbf{x} : 0 \leq x_N \leq \sqrt{\sum_{j=1}^{N-1} x_j^2}\right\} \text{ and } H_- \equiv \left\{\mathbf{x} : -\sqrt{\sum_{j=1}^{N-1} x_j^2} \leq x_N \leq 0\right\},$$

and so, by Theorem 5.3.3 and Corollary 5.3.2, $\overline{B(\mathbf{0}, r)}$ is Riemann measurable. Further, by (5.1.2) and (5.1.1), for any $\mathbf{c} \in \mathbb{R}^N$, $\overline{B(\mathbf{c}, r)}$ is Riemann measurable and $\text{vol}\big(\overline{B(\mathbf{c}, r)}\big) = \text{vol}\big(\overline{B(\mathbf{0}, r)}\big)\Omega_N r^N$, where Ω_N is the volume of the closed unit ball $\overline{B(\mathbf{0}, 1)}$ in $\mathbb{R}^N$.

Proceeding in precisely the same way as we did in the derivation of (5.4.2) and using the identity $b^N - a^N = (b-a)\sum_{k=0}^{N-1} a^k b^{N-1-k}$, we see that, for any continuous $\tilde{f} : [0, r] \longrightarrow \mathbb{C}$,

$$\int_{\overline{B(\mathbf{0},r)}} \tilde{f}(|\mathbf{x}|)\,d\mathbf{x} = \lim_{n\to\infty} \frac{N\Omega_N r}{n} \lim_{n\to\infty} \sum_{m=1}^{n} \tilde{f}(\xi_{m,n})\xi_{m,n}^{N-1},$$

where

$$\xi_{m,n} = \frac{r}{n}\left(\frac{1}{N}\sum_{k=0}^{N-1} m^k (m-1)^{N-1-k}\right)^{\frac{1}{N-1}} \in \left[\tfrac{(m-1)r}{n}, \tfrac{mr}{n}\right],$$

and conclude from this that

$$\int_{\overline{B(\mathbf{0},r)}} \tilde{f}(|\mathbf{x}|)\,d\mathbf{x} = N\Omega_N \int_0^r \tilde{f}(\rho)\rho^{N-1}\,d\rho. \tag{5.4.5}$$

Combining (5.4.5) with (5.4.3), we get an expression for Ω_N. By the same reasoning as we used to derive (5.4.3), one finds that

$$(2\pi)^{\frac{N}{2}} = \left(\int_{\mathbb{R}} e^{-\frac{x^2}{2}}\,dx\right)^N = \lim_{r\to\infty}\int_{B(\mathbf{0},r)} e^{-\frac{|\mathbf{x}|^2}{2}}\,d\mathbf{x} = N\Omega_N \int_0^\infty \rho^{N-1}e^{-\frac{\rho^2}{2}}\,d\rho.$$

Next make the change of variables $\rho = (2t)^{\frac{1}{2}}$ to see that

$$\int_{[0,\infty)} \rho^{N-1}e^{-\frac{\rho^2}{2}}\,d\rho = 2^{\frac{N}{2}-1}\int_0^\infty t^{\frac{N}{2}-1}e^{-t}\,dt = 2^{\frac{N}{2}-1}\Gamma\big(\tfrac{N}{2}\big).$$

Thus, we now know that

$$\Omega_N = \frac{2\pi^{\frac{N}{2}}}{N\Gamma\big(\frac{N}{2}\big)} = \frac{\pi^{\frac{N}{2}}}{\Gamma\big(\frac{N}{2}+1\big)}.$$

By (5.4.4) and induction on N,

$$\Gamma\big(\tfrac{2N+1}{2}\big) = \pi^{\frac{1}{2}}2^{-N}\prod_{k=1}^{N}(2k-1) = \pi^{\frac{1}{2}}\frac{(2N)!}{4^N N!},$$

and therefore

$$\Omega_{2N} = \frac{\pi^N}{N!} \quad \text{and} \quad \Omega_{2N-1} = \frac{4^N\pi^{N-1}N!}{(2N)!} \quad \text{for } N \ge 1.$$

Applying (3.2.4), we find that

$$\Omega_{2N} \sim (2\pi N)^{-\frac{1}{2}}\left(\frac{\pi e}{N}\right)^N \quad \text{and} \quad \Omega_{2N-1} \sim (\sqrt{2}\pi)^{-1}\left(\frac{\pi e}{N}\right)^N.$$

Thus, as N gets large, Ω_N, the volume of the unit ball in $\mathbb{R}^N$, is tending to 0 at a very fast rate. Seeing as the volume of the cube $[-1,1]^N$ that circumscribes $\overline{B(\mathbf{0},1)}$ has volume 2^N, this means that the 2^N corners of $[-1,1]^N$ that are not in $\overline{B(\mathbf{0},1)}$ take up the lion's share of the available space. Hence, if we lived in a large dimensional universe, the biblical tradition that a farmer leave the corners of his field to be harvested by the poor would be very generous.

5.5 Rotation Invariance of Integrals

Because they fit together so nicely, thus far we have been dealing exclusively with rectangles whose sides are parallel to the standard coordinate axes. However, this restriction obscures a basic property of integrals, the property of *rotation invariance*. To formulate this property, recall that $(\mathbf{e}_1,\dots,\mathbf{e}_N) \in (\mathbb{R}^N)^N$ is called an *orthonormal basis* in $\mathbb{R}^N$ if $(\mathbf{e}_i,\mathbf{e}_j)_{\mathbb{R}^N} = \delta_{i,j}$. The standard orthonormal basis $(\mathbf{e}_1^0,\dots,\mathbf{e}_N^0)$ is the one for which $(\mathbf{e}_i^0)_j = \delta_{i,j}$, but there are many others. For example, in $\mathbb{R}^2$, for each $\theta \in [0,2\pi)$, $\big((\cos\theta,\sin\theta),(\mp\sin\theta,\pm\cos\theta)\big)$ is an orthonormal basis, and every orthonormal basis in $\mathbb{R}^2$ is one of these.

A rotation[4] in $\mathbb{R}^N$ is a map $\mathfrak{R}:\mathbb{R}^N \longrightarrow \mathbb{R}^N$ of the form $\mathfrak{R}(\mathbf{x}) = \sum_{j=1}^N x_j\mathbf{e}_j$ where $(\mathbf{e}_1,\dots,\mathbf{e}_N)$ is an orthonormal basis. Obviously $\mathfrak{R}$ is linear in the sense that

[4]The terminology that I am using here is slightly inaccurate. The term *rotation* should be reserved for $\mathfrak{R}$'s for which the determinant of the matrix $\big(((\mathbf{e}_i,\mathbf{e}_j^0)_{\mathbb{R}^N})\big)$ is 1, and I have not made a distinction

$$\mathfrak{R}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathfrak{R}(\mathbf{x}) + \beta\mathfrak{R}(\mathbf{y}).$$

In addition, $\mathfrak{R}$ preserves inner products: $\big(\mathfrak{R}(\mathbf{x}), \mathfrak{R}(\mathbf{y})\big)_{\mathbb{R}^N} = (\mathbf{x}, \mathbf{y})_{\mathbb{R}^N}$. To check this, simply note that

$$\big(\mathfrak{R}(\mathbf{x}), \mathfrak{R}(\mathbf{y})\big)_{\mathbb{R}^N} = \sum_{i,j=1}^N x_i y_j (\mathbf{e}_i, \mathbf{e}_j)_{\mathbb{R}^N} = \sum_{i=1}^N x_i y_i = (\mathbf{x}, \mathbf{y})_{\mathbb{R}^N}.$$

In particular, $|\mathfrak{R}(\mathbf{y}) - \mathfrak{R}(\mathbf{x})| = |\mathbf{y} - \mathbf{x}|$, and so it is clear that $\mathfrak{R}$ is one-to-one and continuous. Further, if $\mathfrak{R}$ and $\mathfrak{R}'$ are rotations, then so is $\mathfrak{R}' \circ \mathfrak{R}$. Indeed, if $(\mathbf{e}_1, \dots, \mathbf{e}_N)$ is the orthonormal basis for $\mathfrak{R}$, then

$$\mathfrak{R}' \circ \mathfrak{R}(\mathbf{x}) = \sum_{j=1}^N x_j \mathfrak{R}'(\mathbf{e}_j),$$

and, since

$$\big(\mathfrak{R}'(\mathbf{e}_i), \mathfrak{R}'(\mathbf{e}_j)\big)_{\mathbb{R}^N} = (\mathbf{e}_i, \mathbf{e}_j)_{\mathbb{R}^N} = \delta_{i,j},$$

$\big(\mathfrak{R}'(\mathbf{e}_1), \dots, \mathfrak{R}'(\mathbf{e}_N)\big)$ is an orthonormal basis. Finally, if $\mathfrak{R}$ is a rotation, then there is a unique rotation $\mathfrak{R}^{-1}$ such that $\mathfrak{R} \circ \mathfrak{R}^{-1} = \mathbf{I} = \mathfrak{R}^{-1} \circ \mathfrak{R}$, where $\mathbf{I}$ is the identity map: $\mathbf{I}(\mathbf{x}) = \mathbf{x}$. To see this, let $(\mathbf{e}_1, \dots, \mathbf{e}_N)$ be the orthonormal bases for $\mathfrak{R}$, and set $\tilde{\mathbf{e}}_i = \big((\mathbf{e}_1)_i, \dots, (\mathbf{e}_N)_i\big)$ for $1 \le i \le N$. Using $(\mathbf{e}_1^0, \dots, \mathbf{e}_N^0)$ to denote the standard orthonormal basis, we have that $(\tilde{\mathbf{e}}_i, \tilde{\mathbf{e}}_j)_{\mathbb{R}^N}$ equals

$$\sum_{k=1}^N (\mathbf{e}_k)_i (\mathbf{e}_k)_j = \sum_{k=1}^N (\mathbf{e}_k, \mathbf{e}_i^0)_{\mathbb{R}^N} (\mathbf{e}_k, \mathbf{e}_j^0)_{\mathbb{R}^N} = \big(\mathfrak{R}(\mathbf{e}_i^0), \mathfrak{R}(\mathbf{e}_j^0)\big)_{\mathbb{R}^N} = \delta_{i,j},$$

and so $(\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_N)$ is an orthonormal basis. Moreover, if $\tilde{\mathfrak{R}}$ is the corresponding rotation, then

$$\begin{aligned}\tilde{\mathfrak{R}} \circ \mathfrak{R}(\mathbf{x}) &= \sum_{i=1}^N x_i \tilde{\mathfrak{R}}(\mathbf{e}_i) = \sum_{i,j=1}^N x_i (\mathbf{e}_i, \mathbf{e}_j^0)_{\mathbb{R}^N} \tilde{\mathbf{e}}_j \\ &= \sum_{i,j,k=1}^N x_i (\mathbf{e}_i, \mathbf{e}_j^0)_{\mathbb{R}^N} (\mathbf{e}_k, \mathbf{e}_j^0)_{\mathbb{R}^N} \mathbf{e}_k^0 = \sum_{i,k=1}^N x_i (\mathbf{e}_i, \mathbf{e}_k)_{\mathbb{R}^N} \mathbf{e}_k^0 = \mathbf{x}.\end{aligned}$$

A similar computation shows that $\mathfrak{R} \circ \tilde{\mathfrak{R}} = \mathbf{I}$, and so we can take $\mathfrak{R}^{-1} = \tilde{\mathfrak{R}}$.

(Footnote 4 continued)
between them and those for which it is -1. That is, I am calling all orthogonal transformations rotations.

Because $\mathfrak{R}$ preserves lengths, it is clear that $\mathfrak{R}\big(\overline{B(\mathbf{c},r)}\big) = \overline{B\big(\mathfrak{R}(\mathbf{c}),r\big)}$. $\mathfrak{R}$ also takes a rectangle into a rectangle, but unfortunately the image rectangle may no longer have sides parallel to the standard coordinate axes. Instead, they are parallel to the axes for the corresponding orthonormal basis. That is,

$$(*) \qquad \mathfrak{R}\left(\prod_{j=1}^N [a_j,b_j]\right) = \left\{\sum_{j=1}^N x_j\mathbf{e}_j : \mathbf{x} \in \prod_{j=1}^N [a_j,b_j]\right\}.$$

Of course, we should expect that $\operatorname{vol}\left(\mathfrak{R}\left(\prod_{j=1}^N [a_j,b_j]\right)\right) = \prod_{j=1}^N (b_j - a_j)$, but this has to be checked, and for that purpose we will need the following lemma.

Lemma 5.5.1 *Let G be a non-empty, bounded, open subset of $\mathbb{R}^N$, and assume that*

$$\lim_{r\searrow 0} \big|(\partial G)^{(r)}\big|_{\mathrm{i}} = 0$$

where $(\partial G)^{(r)}$ is the set of $\mathbf{y}$ for which there exists an $\mathbf{x} \in \partial G$ such that $|\mathbf{y}-\mathbf{x}| < r$. Then $\bar{G}$ is Riemann measurable and, for each $\epsilon > 0$, there exists a finite set $\mathcal{B}$ of mutually disjoint closed balls $\bar{B} \subseteq G$ such that $\operatorname{vol}(\bar{G}) \le \sum_{B\in\mathcal{B}} \operatorname{vol}(\bar{B}) + \epsilon$.

Proof First note that ∂G is Riemann negligible and therefore that $\bar{G}$ is Riemann measurable. Next, given a closed cube $Q = \prod_{j=1}^N [c_j - r, c_j + r]$, let $\bar{B}_Q$ be the closed ball $\overline{B\big(\mathbf{c},\frac{r}{2}\big)}$.

For each $n \ge 1$, let $\mathcal{K}_n$ be the collection of closed cubes Q of the form $2^{-n}\mathbf{k} + [0,2^{-n}]^N$, where $\mathbf{k} \in \mathbb{Z}^N$. Obviously, for each n, the cubes in $\mathcal{K}_n$ are non-overlapping and $\mathbb{R}^N = \bigcup_{Q\in\mathcal{K}_n} Q$.

Now choose n_1 so that $|(\partial G)^{(2^{\frac{N}{2}-n_1})}|_{\mathrm{i}} \le \frac{1}{2}\operatorname{vol}(\bar{G})$, and set

$$\mathcal{C}_1 = \{Q \in \mathcal{K}_{n_1} : Q \subseteq G\} \text{ and } \mathcal{C}_1' = \{Q \in \mathcal{K}_{n_1} : Q \cap G \ne \emptyset\}.$$

Then $\bar{G} \subseteq \bigcup_{Q\in\mathcal{C}_1'} Q$, $\bigcup_{Q\in\mathcal{C}_1'\setminus\mathcal{C}_1} Q \subseteq (\partial G)^{(2^{\frac{N}{2}-n_1})}$, and therefore

$$\sum_{Q\in\mathcal{C}_1} |Q| = \sum_{Q\in\mathcal{C}_1'} |Q| - \sum_{Q\in\mathcal{C}_1'\setminus\mathcal{C}_1} |Q| \ge \operatorname{vol}(\bar{G}) - \frac{\operatorname{vol}(\bar{G})}{2} = \frac{\operatorname{vol}(\bar{G})}{2}.$$

Clearly the $\bar{B}_Q$'s for $Q \in \mathcal{C}_1$ are mutually disjoint, closed balls contained in G. Furthermore, $\operatorname{vol}(\bar{B}_Q) = \alpha|Q|$, where $\alpha \equiv 4^{-N}\Omega_N$, and therefore

$$\operatorname{vol}\left(G\setminus \bigcup_{Q\in\mathcal{C}_1} \bar{B}_Q\right) = \operatorname{vol}(G) - \sum_{Q\in\mathcal{C}_1} \operatorname{vol}(\bar{B}_Q) = \operatorname{vol}(G) - \alpha\sum_{Q\in\mathcal{C}_1} |Q| \le \beta\operatorname{vol}(G),$$

where $\beta \equiv 1 - \frac{\alpha}{2}$. Finally, set $\mathcal{B}_1 = \{\bar{B}_Q : Q \in \mathcal{C}_1\}$.

Set $G_1 = G \setminus \bigcup_{\bar{B} \in \mathcal{B}_1} \bar{B}$. Then G_1 is again a non-empty, bounded, open set. Furthermore, since (cf. Exercise 4.1) $\partial G_1 \subseteq \partial G \cup \bigcup_{\bar{B} \in \mathcal{B}_1} \partial \bar{B}$, it is easy to see that $\lim_{r \searrow 0} \big|(\partial G_1)^{(r)}\big|_{\mathrm{i}} = 0$. Hence we can apply the same argument to G_1 and thereby produce a set $\mathcal{B}_2 \supseteq \mathcal{B}_1$ of mutually disjoint, closed balls $\bar{B}$ such that $\bar{B} \subseteq G_1$ for $\bar{B} \in \mathcal{B}_2 \setminus \mathcal{B}_1$ and

$$\operatorname{vol}\left(G \setminus \bigcup_{\bar{B} \in \mathcal{B}_2} \bar{B}\right) = \operatorname{vol}\left(G_1 \setminus \bigcup_{\bar{B} \in \mathcal{B}_2 \setminus \mathcal{B}_1} B_k\right) \le \beta \operatorname{vol}(G_1) \le \beta^2 \operatorname{vol}(G).$$

After m iterations, we produce a collection $\mathcal{B}_m$ of mutually disjoint closed balls $\bar{B} \subseteq G$ such that $\operatorname{vol}\left(G \setminus \bigcup_{\bar{B} \in \mathcal{B}_m} \bar{B}\right) \le \beta^m \operatorname{vol}(G)$. Thus, all that remains is to choose m so that $\beta^m \operatorname{vol}(G) < \epsilon$ and then do m iterations. □

Lemma 5.5.2 *If R is a rectangle and $\mathfrak{R}$ is a rotation, then $\mathfrak{R}(R)$ is Riemann measurable and has the same volume as R.*

Proof It is obvious that $\operatorname{int}(R)$ satisfies the hypotheses of Lemma 5.5.1, and, by using $(*)$, it is easy to check that $\operatorname{int}\big(\mathfrak{R}(R)\big)$ does also.

Next assume that $G \equiv \operatorname{int}(R) \neq \emptyset$. Clearly G satisfies the hypotheses of Lemma 5.5.1, and therefore for each $\epsilon > 0$ we can find a collection $\mathcal{B}$ of mutually disjoint closed balls $\bar{B} \subseteq G$ such that $\sum_{\bar{B} \in \mathcal{B}} \operatorname{vol}(\bar{B}) + \epsilon \ge \operatorname{vol}(\bar{G}) = \operatorname{vol}(R)$. Thus, if $\mathcal{B}' = \{\mathfrak{R}(\bar{B}) : \bar{B} \in \mathcal{B}\}$, then $\mathcal{B}'$ is a collection of mutually disjoint closed balls $\bar{B}' \subseteq \mathfrak{R}(R)$ such that

$$\operatorname{vol}(R) \le \sum_{\bar{B} \in \mathcal{B}} \operatorname{vol}(\bar{B}) + \epsilon = \sum_{\bar{B}' \in \mathcal{B}'} \operatorname{vol}(\bar{B}') + \epsilon \le \operatorname{vol}\big(\mathfrak{R}(R)\big) + \epsilon,$$

and so $\operatorname{vol}\big(\mathfrak{R}(R)\big) \ge |R|$. To prove that this inequality is an equality, apply the same line of reasoning to $G' = \operatorname{int}\big(\mathfrak{R}(R)\big)$ and $\mathfrak{R}^{-1}$ acting on $\mathfrak{R}(R)$, and thereby obtain

$$\operatorname{vol}(R) = \operatorname{vol}\big(\mathfrak{R}^{-1} \circ \mathfrak{R}(R)\big) \ge \operatorname{vol}\big(\mathfrak{R}(R)\big).$$

Finally, if $R = \emptyset$ there is nothing to do. On the other hand, if $R \neq \emptyset$ but $\operatorname{int}(R) = \emptyset$, for each $\epsilon > 0$ let $R(\epsilon)$ be the set of points $\mathbf{y} \in \mathbb{R}^N$ such that $\max_{1 \le j \le N} |y_j - x_j| \le \epsilon$ for some $\mathbf{x} \in R$. Then $R(\epsilon)$ is a closed rectangle with non-empty interior containing R, and so $\operatorname{vol}\big(\mathfrak{R}(R)\big) \le \operatorname{vol}\big(\mathfrak{R}(R(\epsilon))\big) = |R(\epsilon)|$. Since $\operatorname{vol}(R) = 0 = \lim_{\epsilon \searrow 0} |R(\epsilon)|$, it follows that $\operatorname{vol}(R) = \operatorname{vol}\big(\mathfrak{R}(R)\big)$ in this case also. □

Theorem 5.5.3 *If Γ is a bounded, Riemann measurable subset and $\mathfrak{R}$ is a rotation, then $\mathfrak{R}(\Gamma)$ is Riemann measurable and* $\operatorname{vol}\big(\mathfrak{R}(\Gamma)\big) = \operatorname{vol}(\Gamma)$.

Proof Given $\epsilon > 0$, choose $\mathcal{C}_1$ to be a collection of non-overlapping rectangles contained in Γ such that $\mathrm{vol}(\Gamma) \le \sum_{R\in\mathcal{C}_1} |R| + \epsilon$, and choose $\mathcal{C}_2$ to be a cover of Γ by non-overlapping rectangles such that $\mathrm{vol}(\Gamma) \ge \sum_{R\in\mathcal{C}_2} |R| - \epsilon$. Then

$$\begin{aligned}|\mathfrak{R}(\Gamma)|_{\mathrm{i}} &\ge \sum_{R\in\mathcal{C}_1} \mathrm{vol}\big(\mathfrak{R}(R)\big) = \sum_{R\in\mathcal{C}_1} |R| \ge \mathrm{vol}(\Gamma) - \epsilon \ge \sum_{R\in\mathcal{C}_2} |R| - 2\epsilon \\ &= \sum_{R\in\mathcal{C}_2} \mathrm{vol}\big(\mathfrak{R}(R)\big) - 2\epsilon \ge |\mathfrak{R}(\Gamma)|_{\mathrm{e}} - 2\epsilon.\end{aligned}$$

Hence, $|\mathfrak{R}(\Gamma)|_{\mathrm{e}} \le \mathrm{vol}(\Gamma) + 2\epsilon$ and $|\mathfrak{R}(\Gamma)|_{\mathrm{i}} \ge \mathrm{vol}(\Gamma) - 2\epsilon$ for all $\epsilon > 0$. □

Corollary 5.5.4 *Let $f : \overline{B(0, r)} \longrightarrow \mathbb{C}$ be a bounded function that is continuous off of a Riemann negligible set. Then, for each rotation $\mathfrak{R}$, $f \circ \mathfrak{R}$ is continuous off of a Riemann negligible set and*

$$\int_{\overline{B(\mathbf{0},r)}} f \circ \mathfrak{R}(\mathbf{x})\, d\mathbf{x} = \int_{\overline{B(\mathbf{0},r)}} f(\mathbf{x})\, d\mathbf{x}.$$

Proof Without loss in generality, we will assume throughout that f is real-valued.

If D is a Riemann negligible set off of which f is continuous, then $\mathfrak{R}^{-1}(D)$ contains the set where $f\circ\mathfrak{R}$ is discontinuous. Hence, since $\mathrm{vol}\big(\mathfrak{R}^{-1}(D)\big) = \mathrm{vol}(D) = 0$, $f \circ \mathfrak{R}$ is continuous off of a Riemann negligible set.

Set $g = \mathbf{1}_{\overline{B(\mathbf{0},r)}} f$. Then, by the preceding, both g and $g \circ \mathfrak{R}$ are Riemann integrable. By (5.4.1), for any cover $\mathcal{C}$ of $[-r, r]^N$ by non-overlapping rectangles and any associated choice function $\boldsymbol{\Xi}$,

$$\begin{aligned}\int_{\overline{B(\mathbf{0},r)}} f \circ \mathfrak{R}(\mathbf{x})\, d\mathbf{x} &= \sum_{R\in\mathcal{C}} \int_{\mathfrak{R}^{-1}(R)} g \circ \mathfrak{R}(\mathbf{x})\, d\mathbf{x} \\ &= \mathcal{R}(g; \mathcal{C}, \boldsymbol{\Xi}) + \sum_{R\in\mathcal{C}} \int_{\mathfrak{R}^{-1}(R)} \Delta_R(\mathbf{x})\, d\mathbf{x},\end{aligned}$$

where $\Delta_R(\mathbf{x}) = g(\mathbf{x}) - g\big(\boldsymbol{\Xi}(R)\big)$. Since $\mathcal{R}\big(g; \mathcal{C}, \boldsymbol{\Xi}\big)$ tends to $\int_{\overline{B(\mathbf{0},r)}} f(\mathbf{x})\, d\mathbf{x}$ as $\|\mathcal{C}\| \to 0$, what remains to be shown is that the final term tends to 0 as $\|\mathcal{C}\| \to 0$. But $|\Delta_R(\mathbf{x})| \le \sup_R g - \inf_R g$ and therefore

$$\left|\sum_{R\in\mathcal{C}} \int_{\mathfrak{R}^{-1}(R)} \Delta_R(\mathbf{x})\, d\mathbf{x}\right| \le \sum_{R\in\mathcal{C}} \left(\sup_R g - \inf_R g\right) |R| = \mathcal{U}(g; \mathcal{C}) - \mathcal{L}(g; \mathcal{C}),$$

which tends to 0 as $\|\mathcal{C}\| \to 0$. □

Here is an example of the way in which one can use rotation invariance to make computations.

Lemma 5.5.5 *Let* $0 \le r_1 < r_2$ *and* $0 \le \theta_1 < \theta_2 < 2\pi$ *be given. Then the region*

$$\big\{(r\cos\theta, r\sin\theta) : (r,\theta) \in [r_1, r_2] \times [\theta_1, \theta_2]\big\}$$

has a Riemann negligible boundary and volume $\frac{r_2^2 - r_1^2}{2}(\theta_2 - \theta_1)$.

Proof Because this region can be constructed by taking the intersection of differences of balls with half spaces, its boundary is Riemann negligible. Furthermore, to compute its volume, it suffices to treat the case when $r_1 = 0$ and $r_2 = 1$, since the general case can be reduced to this one by taking differences and scaling.

Now define $u(\theta) = \mathrm{vol}\big(W(\theta)\big)$ where

$$W(\theta) \equiv \big\{(r\cos\omega, r\sin\omega) : (r,\omega) \in [0,1] \times [0,\theta]\big\}.$$

Obviously, u is a non-decreasing function of $\theta \in [0, 2\pi]$ that is equal to 0 when $\theta = 0$ and π when $\theta = 2\pi$. In addition, $u(\theta_1 + \theta_2) = u(\theta_1) + u(\theta_2)$ if $\theta_1 + \theta_2 \le 2\pi$. To see this, let $\mathfrak{R}_{\theta_1}$ be the rotation corresponding to the orthonormal basis $\big((\cos\theta_1, \sin\theta_1), (-\sin\theta_1, \cos\theta_1)\big)$, and observe that

$$W(\theta_1 + \theta_2) = W(\theta_1) \cup \mathfrak{R}_{\theta_1}\big(W(\theta_2)\big)$$

and that $\mathrm{int}\big(W(\theta_1)\big) \cap \mathrm{int}\big(\mathfrak{R}_{\theta_1}\big(W(\theta_2)\big)\big) = \emptyset$. Hence, the equality follows from the facts that the boundaries of $W(\theta_1)$ and $\mathfrak{R}\big(W(\theta_2)\big)$ are Riemann negligible and that $\mathfrak{R}_{\theta_1}\big(W(\theta_2)\big)$ has the same volume as $W(\theta_2)$. After applying this repeatedly, we get $nu\big(\frac{2\pi}{n}\big) = \pi$ and then that $u\big(\frac{2\pi m}{n}\big) = mu\big(\frac{2\pi}{n}\big)$ for $n \ge 1$ and $0 \le m \le n$. Hence, $u\big(\frac{2\pi m}{n}\big) = \frac{\pi m}{n}$ for all $n \ge 1$ and $0 \le m \le n$. Now, given any $\theta \in (0, 2\pi)$, choose $\{m_n \in \mathbb{N} : n \ge 1\}$ so that $0 \le \theta - \frac{2\pi m_n}{n} < \frac{2\pi}{n}$. Then, for all $n \ge 1$,

$$\big|u(\theta) - \tfrac{\theta}{2}\big| \le \big|u(\theta) - u\big(\tfrac{2\pi m_n}{n}\big)\big| + \big|\tfrac{\pi m_n}{n} - \tfrac{\theta}{2}\big| \le u\big(\tfrac{2\pi}{n}\big) + \tfrac{\pi}{n} \le \tfrac{2\pi}{n},$$

and so $u(\theta) = \frac{\theta}{2}$.

Finally, given any $0 \le \theta_1 < \theta_2 \le 2\pi$, set $\theta = \theta_2 - \theta_1$, and observe that $W(\theta_2)\backslash\mathrm{int}(W(\theta_1) = \mathfrak{R}_{\theta_1}\big(W(\theta)\big)$ and therefore that $W(\theta_2)\backslash\mathrm{int}(W(\theta_1)$ has the same volume as $W(\theta)$. □

5.6 Polar and Cylindrical Coordinates

Changing variables in multi-dimensional integrals is more complicated than in one dimension. From the standpoint of the theory that we have developed, the primary reason is that, in general, even linear changes of coordinates take rectangles into parallelograms that are not in general rectangles with respect to any orthonormal basis. Starting from the formula in terms of determinants for the volume of parallelograms, Jacobi worked out a general formula that says how integrals transform under continuously differentiable changes that satisfy a suitable non-degeneracy condition, but his theory relies on a familiarity with quite a lot of linear algebra and matrix theory. Thus, we will restrict our attention to changes of variables for which his general theory is not required.

We will begin with *polar coordinates* for $\mathbb{R}^2$. To every point $\mathbf{x} \in \mathbb{R}^2\backslash\{\mathbf{0}\}$ there exists a unique point $(\rho, \varphi) \in (0, \infty) \times [0, 2\pi)$ such that $x_1 = \rho\cos\varphi$ and $x_2 = \rho\sin\varphi$. Indeed, if $\rho = |\mathbf{x}|$, then $\frac{\mathbf{x}}{\rho} \in \mathbb{S}^1(\mathbf{0}, 1)$, and so φ is the distance, measured counterclockwise, one travels along $\mathbb{S}^1(\mathbf{0}, 1)$ to get from $(1, 0)$ to $\frac{\mathbf{x}}{\rho}$. Thus we can use the variables $(\rho, \varphi) \in (0, \infty) \times [0, 2\pi)$ to parameterize $\mathbb{R}^2\backslash\{\mathbf{0}\}$. We have restricted our attention to $\mathbb{R}^2\backslash\{\mathbf{0}\}$ because this parameterization breaks down at $\mathbf{0}$. Namely, $\mathbf{0} = (0\cos\varphi, 0\sin\varphi)$ for *every* $\varphi \in [0, 2\pi)$. However, this flaw will not cause us problems here.

Given a continuous function $f : \overline{B(\mathbf{0}, r)} \longrightarrow \mathbb{C}$, it is reasonable to ask whether the integral of f over $\overline{B(\mathbf{0}, r)}$ can be written as an integral with respect to the variables (ρ, φ). In fact, we have already seen in (5.4.2) that this is possible when f depends only on $|\mathbf{x}|$, and we will now show that it is always possible. To this end, for $\theta \in \mathbb{R}$, let $\mathfrak{R}_\theta$ be the rotation in $\mathbb{R}^2$ corresponding to the basis $\big((\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\big)$. That is,

$$\mathfrak{R}_\theta\mathbf{x} = \big(x_1\cos\theta - x_2\sin\theta, x_1\sin\theta + x_2\cos\theta\big).$$

Using (1.5.1), it is easy to check that $\mathfrak{R}_\theta \circ \mathfrak{R}_\varphi = \mathfrak{R}_{\theta+\varphi}$. In particular, $\mathfrak{R}_{2\pi+\varphi} = \mathfrak{R}_\varphi$.

Lemma 5.6.1 *Let $f : \overline{B(\mathbf{0}, r)} \longrightarrow \mathbb{C}$ be a continuous function, and define*

$$\tilde{f}(\rho) = \frac{1}{2\pi}\int_0^{2\pi} f\big(\rho\cos\varphi, \rho\sin\varphi\big)\, d\varphi \quad \textit{for } \rho \in [0, r].$$

Then, for all $\mathbf{x} \in \overline{B(\mathbf{0}, r)}$,

$$\tilde{f}(|\mathbf{x}|) = \frac{1}{2\pi}\int_0^{2\pi} f\big(\mathfrak{R}_\varphi\mathbf{x}\big)\, d\varphi.$$

Proof Set $\rho = |\mathbf{x}|$ and choose $\theta \in [0, 1)$ so that $\mathbf{x} = \big(\rho\cos(2\pi\theta), \rho\sin(2\pi\theta)\big)$. Equivalently, $\mathbf{x} = \mathfrak{R}_{2\pi\theta}(\rho, 0)$. Then by the preceding remarks about rotations in $\mathbb{R}^2$ and (3.3.3) applied to the periodic function $\xi \rightsquigarrow f\big(\mathfrak{R}_{2\pi\xi}(\rho, 0)\big)$,

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} f\big(\mathfrak{R}_\varphi\mathbf{x}\big)\, d\varphi &= \frac{1}{2\pi}\int_0^{2\pi} f\big(\mathfrak{R}_{2\pi\theta+\varphi}(\rho, 0)\big) \\ &= \int_0^1 f\big(\mathfrak{R}_{2\pi(\theta+\varphi)}(\rho, 0)\big)\, d\varphi = \int_0^1 f\big(\mathfrak{R}_{2\pi\varphi}(\rho, 0)\big)\, d\varphi = \tilde{f}(\rho).\end{aligned}$$

□

Theorem 5.6.2 *If f is a continuous function on the ball $\overline{B(\mathbf{0},r)}$ in $\mathbb{R}^2$, then*

$$\int_{\overline{B(\mathbf{0},r)}} f(\mathbf{x})\,d\mathbf{x} = \int_0^r \rho\left(\int_0^{2\pi} f\big(\rho\cos\varphi, \rho\sin\varphi\big)\,d\varphi\right)d\rho$$
$$= \int_0^{2\pi}\left(\int_0^r f\big(\rho\cos\varphi, \rho\sin\varphi\big)\rho\,d\rho\right)d\varphi.$$

Proof By (5.4.2),

$$\int_{\overline{B(\mathbf{0},r)}} f(\mathbf{x})\,d\mathbf{x} = \int_{\overline{B(\mathbf{0},r)}} f\big(\mathfrak{R}_\varphi\mathbf{x}\big)\,d\mathbf{x}$$

for all φ. Hence, by (5.2.1), Lemma 5.6.1, and (5.4.2),

$$\int_{\overline{B(\mathbf{0},r)}} f(\mathbf{x})\,d\mathbf{x} = \frac{1}{2\pi}\int_0^{2\pi}\left(\int_{\overline{B(\mathbf{0},r)}} f\big(\mathfrak{R}_\varphi\mathbf{x}\big)\,d\mathbf{x}\right)d\varphi$$
$$= \int_{\overline{B(\mathbf{0},r)}} \tilde{f}(|\mathbf{x}|)\,d\mathbf{x} = \int_0^r \tilde{f}(\rho)\rho\,d\rho,$$

which is the first equality. The second equality follows from the first by another application of (5.2.1). □

As a preliminary application of this theorem, we will use it to compute integrals over a *star shaped region*, a region G for which there exists a $\mathbf{c} \in \mathbb{R}^2$, known as the *center*, and a continuous function, known as the *radial function*, $r : [0, 2\pi] \longrightarrow (0,\infty)$ such that $r(0) = r(2\pi)$ and

$$G = \big\{\mathbf{c} + r\mathbf{e}(\varphi) : \varphi \in [0, 2\pi)\ \&\ r \in \big[0, r(\varphi)\big)\big\}, \tag{5.6.1}$$

where $\mathbf{e}(\varphi) \equiv (\cos\varphi, \sin\varphi)$. For instance, if G is a non-empty, bounded, convex open set, then for any $\mathbf{c} \in G$, G is star shaped with center at $\mathbf{c}$ and

$$r(\varphi) = \max\{r > 0 : \mathbf{c} + r\mathbf{e}(\varphi) \in \bar{G}\}.$$

Observe that

$$\partial G = \big\{\mathbf{c} + r(\varphi)\mathbf{e}(\varphi) : \varphi \in [0, 2\pi)\big\}.$$

and, as a consequence, we can show that ∂G is Riemann negligible. Indeed, for a given $\epsilon \in (0, 1]$ choose $n \geq 1$ so that $|r(\varphi_2) - r(\varphi_1)| < \epsilon$ if $|\varphi_2 - \varphi_1| \leq \frac{2\pi}{n}$ and, for $1 \leq m \leq n$, set

$$A_m = \big\{\mathbf{c} + \rho\mathbf{e}(\varphi) : \tfrac{2\pi(m-1)}{n} \leq \varphi < \tfrac{2\pi m}{n}\ \&\ \big|\rho - r\big(\tfrac{2\pi m}{n}\big)\big| \leq \epsilon\big\}.$$

Then $\partial G \subseteq \bigcup_{m=1}^{n} A_{m,n}$ and, by Lemma 5.5.5,

$$\mathrm{vol}(A_m) = \frac{2\pi^2}{n}\Big(\big(r\big(\tfrac{2\pi m}{n}\big)+\epsilon\big)^2 - \big(r\big(\tfrac{2\pi m}{n}\big)-\epsilon\big)^2\Big) \le \frac{8\pi^2\|r\|_{[0,2\pi]}\epsilon}{n},$$

and therefore there is a constant $K < \infty$ such that $|\partial G|_{\mathrm{e}} \le K\epsilon$ for all $\epsilon \in (0, 1]$. Finally, notice that G is path connected and therefore, by Exercise 4.5 is connected.

The following is a significant extension of Theorem 5.6.2.

Corollary 5.6.3 *If G is the region in (5.6.1) and $f : \bar{G} \longrightarrow \mathbb{C}$ is continuous, then*

$$\int_{\bar{G}} f(\mathbf{x})\,d\mathbf{x} = \int_0^{2\pi}\left(\int_0^{r(\varphi)} f(\mathbf{c}+\rho\mathbf{e}(\varphi))\,\rho d\rho\right)d\varphi.$$

Proof Without loss in generality, we will assume that $\mathbf{c} = \mathbf{0}$. Set $r_- = \min\{r(\varphi) : \varphi \in [0, 2\pi]\}$ and $r_+ = \max\{r(\varphi) : \varphi \in [0, 2\pi]\}$. Given $n \ge 1$, define $\eta_n : \mathbb{R} \longrightarrow [0, 1]$ by

$$\eta_n(t) = \begin{cases} 0 & \text{if } t \le 0 \\ \frac{nt}{r_-} & \text{if } 0 < t \le \frac{r_-}{n} \\ 1 & \text{if } t > \frac{r_-}{n}, \end{cases}$$

and define α_n and β_n on $\mathbb{R}^2$ by

$$\alpha_n(\rho\mathbf{e}(\varphi)) = \eta_n\big(r(\varphi)-\rho\big) \text{ and } \beta_n(\rho\mathbf{e}(\varphi)) = \eta_n\big(r(\varphi)+\tfrac{r_-}{n}-\rho\big)$$

Then both α_n and β_n are continuous functions, α_n vanishes off of G and β_n equals 1 on $\bar{G}$. Finally define

$$f_n(\mathbf{x}) \equiv \begin{cases} \alpha_n(\mathbf{x})f(\mathbf{x}) & \text{if } \mathbf{x} \in G \\ 0 & \text{if } \mathbf{x} \notin G. \end{cases}$$

Then f_n is continuous and therefore, by Theorem 5.6.2,

$$\begin{aligned}\int_{\bar{G}} f_n(\mathbf{x})\,d\mathbf{x} &= \int_{\overline{B(\mathbf{0},r_+)}} f_n(\mathbf{x})\,d\mathbf{x} = \int_0^{2\pi}\left(\int_0^{r_+} f_n(\rho\mathbf{e}(\varphi))\rho\,d\rho\right)d\varphi \\ &= \int_0^{2\pi}\left(\int_0^{r(\varphi)} f_n(\rho\mathbf{e}(\varphi))\rho\,d\rho\right)d\varphi.\end{aligned}$$

Clearly, again by Theorem 5.6.2,

$$\left|\int_{\bar{G}} f(\mathbf{x})\,d\mathbf{x} - \int_{\bar{G}} f_n(\mathbf{x})\,d\mathbf{x}\right| \le \|f\|_{\bar{G}} \int_{\bar{G}} \big(1-\alpha_n(\mathbf{x})\big)\,d\mathbf{x}$$
$$\le \|f\|_{\bar{G}} \int_{\overline{B(\mathbf{0},2r_+)}} \beta_n(\mathbf{x})\big(1-\alpha_n(\mathbf{x})\big)\,d\mathbf{x}$$
$$= \|f\|_{\bar{G}} \int_0^{2\pi} \left(\int_0^{2r_+} \eta_n\big(r(\varphi)+\tfrac{r_-}{n}-\rho\big)\Big(1-\eta_n\big(r(\varphi)-\rho\big)\Big)\rho\,d\rho\right)d\varphi$$
$$= \|f\|_{\bar{G}} \int_0^{2\pi} \left(\int_{r(\varphi)-\frac{r_-}{n}}^{r(\varphi)+\frac{r_-}{n}} \eta_n\big(r(\varphi)+\tfrac{r_-}{n}-\rho\big)\Big(1-\eta_n\big(r(\varphi)-\rho\big)\Big)\rho\,d\rho\right)d\varphi$$
$$\le \frac{8\pi\|f\|_{\bar{G}}r_+r_-}{n}.$$

At the same time,

$$\left|\int_0^{2\pi}\left(\int_0^{r(\varphi)} f(\rho\mathbf{e}(\varphi))\rho\,d\rho\right)d\varphi - \int_0^{2\pi}\left(\int_0^{r(\varphi)} f_n(\rho\mathbf{e}(\varphi))\rho\,d\rho\right)d\varphi\right|$$
$$\le \|f\|_{\bar{G}}\int_0^{2\pi}\left(\int_{r(\varphi)-\frac{r_-}{n}}^{r(\varphi)} \rho\,d\rho\right)d\varphi \le \frac{4\pi\|f\|_{\bar{G}}r_+r_-}{n}.$$

Thus, the asserted equality follows after one lets $n\to\infty$. □

We turn next to *cylindrical coordinates* in $\mathbb{R}^3$. That is, we represent points in $\mathbb{R}^3$ as $(\rho\mathbf{e}(\varphi),\xi)$, where $\rho\ge 0$, $\varphi\in[0,2\pi)$, and $\xi\in\mathbb{R}$. Again the correspondence fails to be one-to-one everywhere. Namely, φ is not uniquely determined for $\mathbf{x}\in\mathbb{R}^3$ with $x_1=x_2=0$, but, as before, this will not prevent us from representing integrals in terms of the variables (ρ,φ,ξ).

Theorem 5.6.4 *Let $\psi:[a,b]\longrightarrow[0,\infty)$ be a continuous function, and set*

$$\Gamma=\{\mathbf{x}\in\mathbb{R}^3:\ x_3\in[a,b]\ \&\ x_1^2+x_2^2\le\psi(x_3)^2\}.$$

Then Γ is Riemann measurable and

$$\int_\Gamma f(\mathbf{x})\,d\mathbf{x}=\int_a^b\left(\int_0^{\psi(\xi)}\rho\left(\int_0^{2\pi} f\big(\rho\mathbf{e}(\varphi),\xi\big)\,d\varphi\right)d\rho\right)d\xi$$

for any continuous function $f:\Gamma\longrightarrow\mathbb{C}$.

Proof Given $n\ge 1$, define $c_{m,n}=\big(1-\frac{m}{n}\big)a+\frac{m}{n}b$ for $0\le m\le n$, and set $I_{m,n}=[c_{m-1,n},c_{m,n}]$ and $\Gamma_{m,n}=\{\mathbf{x}\in\Gamma:\ x_3\in I_{m,n}\}$ for $1\le m\le n$. Next, for each $1\le m\le n$, set $\kappa_{m,n}=\min_{I_{m,n}}\psi$, $K_{m,n}=\max_{I_{m,n}}\psi$, and

$$D_{m,n}=\{\mathbf{x}:\ \kappa_{m,n}^2\le x_1^2+x_2^2\le K_{m,n}^2\ \&\ x_3\in I_{m,n}\}.$$

To see that Γ is Riemann measurable, we will show that its boundary is Riemann negligible. Indeed, $\partial\Gamma_{m,n} \subseteq D_{m,n}$, and therefore, by Theorem 5.2.1 and Lemma 5.5.5, $|\partial\Gamma_{m,n}|_e \le \frac{\pi(K_{m,n}^2-\kappa_{m,n}^2)(b-a)}{n}$. Since

$$\lim_{n\to\infty}\max_{1\le m\le n}(K_{m,n}-\kappa_{m,n})=0 \text{ and } |\partial\Gamma|_e \le \sum_{m=1}^{n}|\partial\Gamma_{m,n}|_e,$$

it follows that $|\partial\Gamma|_e = 0$. Of course, since each $\Gamma_{m,n}$ is a set of the same form as Γ, each of them is also Riemann measurable.

Now let f be given. Then

$$\int_\Gamma f(\mathbf{x})\,d\mathbf{x} = \sum_{m=1}^{n}\int_{\Gamma_{m,n}} f(\mathbf{x})\,d\mathbf{x} = \sum_{m=1}^{n}\int_{C_{m,n}} f(\mathbf{x})\,d\mathbf{x} - \sum_{m=1}^{n}\int_{\Gamma_{m,n}\setminus C_{m,n}} f(\mathbf{x})\,d\mathbf{x},$$

where $C_{m,n} \equiv \{\mathbf{x} : x_1^2+x_2^2 \le \kappa_{m,n} \;\&\; x_3 \in I_{m,n}\}$. Since $\Gamma_{m,n}\setminus C_{m,n} \subseteq D_{m,n}$, the computation in the preceding paragraph shows that

$$\left|\sum_{m=1}^{n}\int_{\Gamma_{m,n}\setminus C_{m,n}} f(\mathbf{x})\,d\mathbf{x}\right| \le \frac{\|f\|_\Gamma \pi(b-a)}{n}\sum_{m=1}^{n}\left(K_{m,n}^2-\kappa_{m,n}^2\right) \longrightarrow 0$$

as $n\to\infty$. Next choose $\xi_{m,n}\in I_{m,n}$ so that $\psi(\xi_{m,n})=\kappa_{m,n}$, and set

$$\epsilon_n = \max_{1\le m\le n}\sup_{\mathbf{x}\in C_{m,n}}|f(\mathbf{x})-f(x_1,x_2,\xi_{m,n})|.$$

Then

$$\left|\sum_{m=1}^{n}\int_{C_{m,n}} f(\mathbf{x})\,d\mathbf{x} - \sum_{m=1}^{n}\int_{C_{m,n}} f(x_1,x_2,\xi_{m,n})\,d\mathbf{x}\right| \le \epsilon_n \text{vol}(\Gamma) \longrightarrow 0.$$

Finally, observe that $\int_{C_{m,n}} f(x_1,x_2,\xi_{m,n})\,d\mathbf{x} = \frac{b-a}{n}g(\xi_{m,n})$ where g is the continuous function on $[a,b]$ given by

$$g(\xi) \equiv \int_0^{\psi(\xi)}\rho\left(\int_0^{2\pi} f\left(\rho\mathbf{e}(\varphi),\xi\right)d\varphi\right)d\rho.$$

Hence, $\sum_{m=1}^{n}\int_{C_{m,n}} f(\mathbf{x})\,d\mathbf{x} = \mathcal{R}(g;\mathcal{C}_n,\Xi_n)$ where $\mathcal{C}_n = \{I_{m,n} : 1\le m\le n\}$ and $\Xi_n(I_{m,n}) = \xi_{m,n}$. Now let $n\to\infty$ to get the desired conclusion. □

Integration over balls in $\mathbb{R}^3$ is a particularly important example to which Theorem 5.6.4 applies. Namely, take $a=-r$, $b=r$, and $\psi(\xi)=\sqrt{r^2-\xi^2}$ for $\xi\in[-r,r]$. Then Theorem 5.6.4 says that

$$\int_{\overline{B(\mathbf{0},r)}} f(\mathbf{x})\,d\mathbf{x} = \int_{-r}^{r}\left(\int_0^{\sqrt{r^2-\xi^2}} \rho\left(\int_0^{2\pi} f\big(\rho\cos\varphi, \rho\sin\varphi, \xi\big)\,d\varphi\right)d\rho\right)d\xi. \tag{5.6.2}$$

There is a beautiful application of (5.6.2) to a famous observation made by Newton about his *law of gravitation*. According to his law, the force exerted by a particle of mass m_1 at $\mathbf{y} \in \mathbb{R}^3$ on a particle of mass m_2 at $\mathbf{b} \in \mathbb{R}^3 \setminus \{\mathbf{x}\}$ is equal to

$$\frac{Gm_1m_2}{|\mathbf{y}-\mathbf{b}|^3}(\mathbf{y}-\mathbf{b}),$$

where G is the gravitational constant. Next, suppose that Ω is a bounded, closed, Riemann measurable region on which mass is continuously distributed with density μ. Then the force that the mass in Ω exerts on a particle of mass m at $\mathbf{b} \notin \Omega$ is given by

$$\int_\Omega \frac{Gm\mu(\mathbf{y})}{|\mathbf{y}-\mathbf{b}|^3}(\mathbf{y}-\mathbf{b})\,d\mathbf{y}.$$

Newton's observation was that if Ω is a ball and the mass density depends only on the distance from the center of the ball, then the force felt by a particle outside the ball is the same as the force exerted on it by a particle at the center of the ball with mass equal to the total mass of the ball. That is, if $\Omega = \overline{B(\mathbf{c}, r)}$ and $\mu : [0, r] \longrightarrow [0, \infty)$ is continuous, then for $\mathbf{b} \notin \overline{B(\mathbf{c}, r)}$,

$$\begin{gathered}\int_{\overline{B(\mathbf{c},r)}} \frac{Gm\mu(|\mathbf{y}-\mathbf{c}|)}{|\mathbf{y}-\mathbf{b}|^3}(\mathbf{y}-\mathbf{b})\,d\mathbf{y} = \frac{GMm}{|\mathbf{c}-\mathbf{b}|^3}(\mathbf{c}-\mathbf{b}) \\ \text{where } M = \int_{\overline{B(\mathbf{c},r)}} \mu\big(|\mathbf{y}-\mathbf{c}|\big)\,d\mathbf{y}.\end{gathered} \tag{5.6.3}$$

(See Exercise 5.8 for the case when $\mathbf{b}$ lies inside the ball).

Using translation and rotations, one can reduce the proof of (5.6.3) to the case when $\mathbf{c} = \mathbf{0}$ and $\mathbf{b} = (0, 0, -D)$ for some $D > r$. Further, without loss in generality, we will assume that $Gm = 1$. Next observe that, by rotation invariance applied to the rotations that take (y_1, y_2, y_3) to $(\mp y_1, \pm y_2, y_3)$,

$$\int_{\overline{B(\mathbf{0},r)}} \frac{\mu(|\mathbf{y}|)}{|\mathbf{y}-\mathbf{b}|^3} y_i\,d\mathbf{y} = -\int_{\overline{B(\mathbf{0},r)}} \frac{\mu(|\mathbf{y}|)}{|\mathbf{y}-\mathbf{b}|^3} y_i\,d\mathbf{y}$$

and therefore

$$\int_{\overline{B(\mathbf{0},r)}} \frac{\mu(|\mathbf{y}|)}{|\mathbf{y}-\mathbf{b}|^3} y_i\,d\mathbf{y} = 0 \quad \text{for } i \in \{1, 2\}.$$

Thus, it remains to show that

$$(*) \qquad \int_{B(\mathbf{0},r)} \frac{\mu(|\mathbf{y}|)}{|\mathbf{y}-\mathbf{b}|^3} y_3 \, d\mathbf{y} = D^{-2} \int_{B(\mathbf{0},r)} \mu(|\mathbf{y}|) \, d\mathbf{y}.$$

To prove $(*)$, we apply (5.6.2) to the function

$$f(\mathbf{x}) = \frac{\mu(|\mathbf{x}|)(x_3 + D)}{\big(x_1^2 + x_2^2 + (x_3 + D)^2\big)^{\frac{3}{2}}}$$

to write the left hand side as $2\pi J$ where

$$J \equiv \int_{-r}^{r} \left(\int_0^{\sqrt{r^2-\xi^2}} \frac{\rho\mu(\sqrt{\rho^2+\xi^2})(\xi + D)}{\big(\rho^2 + (\xi + D)^2\big)^{\frac{3}{2}}} \, d\rho \right) d\xi.$$

Now make the change of variables $\sigma = \sqrt{\rho^2 + \xi^2}$ in the inner integral to see that

$$J = \int_{-r}^{r} (\xi + D) \left(\int_{|\xi|}^{r} \frac{\sigma\mu(\sigma)}{(\sigma^2 + 2\xi D + D^2)^{\frac{3}{2}}} \, d\sigma \right) d\xi,$$

and then apply (5.2.1) to obtain

$$J = \int_0^r \sigma\mu(\sigma) \left(\int_{-\sigma}^{\sigma} \frac{D + \xi}{(\sigma^2 + 2\xi D + D^2)^{\frac{3}{2}}} \, d\xi \right) d\sigma.$$

Use the change of variables $\eta = \sigma^2 + 2\xi D + D^2$ in the inner integral to write it as

$$\frac{1}{4D^2} \int_{(D-\sigma)^2}^{(D+\sigma)^2} \big(\eta^{-\frac{1}{2}} + (D^2 - \sigma^2)\eta^{-\frac{3}{2}}\big) \, d\eta = \frac{2\sigma}{D^2}.$$

Hence,

$$2\pi J = \frac{4\pi}{D^2} \int_0^r \mu(\sigma)\sigma^2 \, d\sigma.$$

Finally, note that $3\Omega_3 = 4\pi$, and apply (5.4.5) with $N = 3$ to see that

$$4\pi \int_0^r \mu(\sigma)\sigma^2 \, d\sigma = \int_{B(\mathbf{0},r)} \mu(|\mathbf{x}|) \, d\mathbf{x}.$$

We conclude this section by using (5.6.2) to derive the analog of Theorem 5.6.2 for integrals over balls in $\mathbb{R}^3$. One way to introduce polar coordinates for $\mathbb{R}^3$ is to think about the use of latitude and longitude to locate points on a globe. To begin with, one has to choose a reference axis, which in the case of a globe is chosen to be the one passing through the north and south poles. Given a point $\mathbf{q}$ on the globe, consider a plane $P_{\mathbf{q}}$ containing the reference axis that passes through $\mathbf{q}$. (There will be only one unless $\mathbf{q}$ is a pole.) Thinking of points on the globe as vectors with base at the center of the earth, the latitude of a point is the angle that $\mathbf{q}$ makes in $P_{\mathbf{q}}$ with the north pole $\mathbf{N}$. Before describing the longitude of $\mathbf{q}$, one has to choose a reference point $\mathbf{q}_0$ that is not on the reference axis. In the case of a globe, the standard choice is Greenwich, England. Then the longitude of $\mathbf{q}$ is the angle between the projections of $\mathbf{q}$ and $\mathbf{q}_0$ in the equatorial plane, the plane that passes through the center of the earth and is perpendicular to the reference axis.

Now let $\mathbf{x} \in \mathbb{R}^3\setminus\{\mathbf{0}\}$. With the preceding in mind, we say that the *polar angle* of $\mathbf{x} = (x_1, x_2, x_3)$ is the $\theta \in [0, \pi]$ such that $\cos\theta = \frac{(\mathbf{x},\mathbf{N})_{\mathbb{R}^3}}{|\mathbf{x}|}$, where $\mathbf{N} = (0, 0, 1)$. Assuming that $\sigma = \sqrt{x_1^2 + x_2^2} > 0$, the *azimuthal angle* of $\mathbf{x}$ is the $\varphi \in [0, 2\pi)$ such that $(x_1, x_2) = (\sigma\cos\varphi, \sigma\sin\varphi)$. In other words, in terms of the globe model, we have taken the center of the earth to lie at the origin, the north pole and south poles to be $(0, 0, 1)$ and $(0, 0, -1)$, and "Greenwich" to be located at $(1, 0, 0)$. Thus the polar angle gives the latitude and the azimuthal angle gives the longitude.

The preceding considerations lead to the parameterization

$$\begin{aligned}&(\rho, \theta, \varphi) \in [0, \infty) \times [0, \pi] \times [0, 2\pi)\\ &\qquad\longmapsto \mathbf{x}_{(\rho,\theta,\varphi)} \equiv \big(\rho\sin\theta\cos\varphi, \rho\sin\theta\sin\varphi, \rho\cos\theta\big) \in \mathbb{R}^3\end{aligned}$$

of points in $\mathbb{R}^3$. Assuming that $\rho > 0$, θ is the polar angle of $\mathbf{x}_{(\rho,\theta,\varphi)}$, and, assuming that $\rho > 0$ and $\theta \notin \{0, \pi\}$, φ is its azimuthal angle. On the other hand, when $\rho = 0$, then $\mathbf{x}_{(\rho,\theta,\varphi)} = \mathbf{0}$ for all $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi)$, and when $\rho > 0$ but $\theta \in \{0, \pi\}$, θ is uniquely determined but φ is not. In spite of these ambiguities, if $\mathbf{x} = \mathbf{x}_{(\rho,\theta,\varphi)}$, then (ρ, θ, φ) are called the *polar coordinates* of $\mathbf{x}$, and as we are about to show, integrals of functions over balls in $\mathbb{R}^3$ can be written as integrals with respect to the variables (ρ, θ, φ).

Let $f : \overline{B(\mathbf{0}, r)} \longrightarrow \mathbb{C}$ be a continuous function. Then, by (5.6.2) and (5.2.2), the integral of f over $\overline{B(\mathbf{0}, r)}$ equals

$$\int_0^{2\pi}\left(\int_{-r}^{r}\left(\int_0^{\sqrt{r^2-\xi^2}} \sigma f_\varphi\big(\sigma, \xi\big)\, d\sigma\right)d\xi\right)d\varphi,$$

where $f_\varphi(\sigma,\xi) = f\big(\sigma\cos\varphi, \sigma\sin\varphi, \xi\big)$. Observe that

$$\begin{aligned}
&\int_{-r}^{r}\left(\int_0^{\sqrt{r^2-\xi^2}} \sigma f_\varphi(\sigma,\xi)\,d\sigma\right)d\xi \\
&\quad = \int_0^{r}\left(\int_0^{\sqrt{r^2-\xi^2}} \sigma f_\varphi(\sigma,\xi)\,d\sigma\right)d\xi + \int_{-r}^{0}\left(\int_0^{\sqrt{r^2-\xi^2}} \sigma f_\varphi(\sigma,\xi)\,d\sigma\right)d\xi \\
&\quad = \int_0^{r} \sigma\left(\int_0^{\sqrt{r^2-\sigma^2}} f_\varphi(\sigma,\xi)\,d\xi\right)d\sigma + \int_0^{r} \sigma\left(\int_0^{\sqrt{r^2-\sigma^2}} f_\varphi(\sigma,-\xi)\,d\xi\right)d\sigma,
\end{aligned}$$

and make the change of variables $\xi = \sqrt{\rho^2-\sigma^2}$ to write

$$\int_0^{\sqrt{r^2-\sigma^2}} f_\varphi(\sigma,\pm\xi)\,d\xi = \int_{(\sigma,r]} f_\varphi\big(\sigma,\pm\sqrt{\rho^2-\sigma^2}\big)\frac{\rho}{\sqrt{\rho^2-\sigma^2}}\,d\rho.$$

Hence, we now know that

$$\begin{aligned}
&\int_{-r}^{r}\left(\int_0^{\sqrt{r^2-\xi^2}} \sigma f_\varphi(\sigma,\xi)\,d\sigma\right)d\xi \\
&\quad = \int_0^{r} \sigma\left(\int_{(\sigma,r]}\Big(f_\varphi\big(\sigma,\sqrt{\rho^2-\sigma^2}\big) + f_\varphi\big(\sigma,-\sqrt{\rho^2-\sigma^2}\big)\Big)\frac{\rho}{\sqrt{\rho^2-\sigma^2}}\,d\rho\right)d\sigma \\
&\quad = \int_0^{r} \rho\left(\int_{[0,\rho)}\Big(f_\varphi\big(\sigma,\sqrt{\rho^2-\sigma^2}\big) + f_\varphi\big(\sigma,-\sqrt{\rho^2-\sigma^2}\big)\Big)\frac{\sigma}{\sqrt{\rho^2-\sigma^2}}d\sigma\right)d\rho,
\end{aligned}$$

where we have made use of the obvious extension of Fubini's Theorem to integrals that are limits of Riemann integrals. Finally, use the change of variables $\sigma = \rho\sin\theta$ in the inner integral to conclude that

$$\int_{-r}^{r}\left(\int_0^{\sqrt{r^2-\xi^2}} \sigma f_\varphi(\sigma,\xi)\,d\sigma\right)d\xi = \int_0^{r} \rho^2\left(\int_0^{\pi} f_\varphi\big(\rho\sin\theta,\rho\cos\theta\big)\,d\theta\right)d\rho$$

and therefore, after an application of (5.2.2), that

$$\begin{aligned}
&\int_{\overline{B(\mathbf{0},r)}} f(\mathbf{x})\,d\mathbf{x} \\
&= \int_0^{r} \rho^2\left(\int_0^{\pi}\left(\int_0^{2\pi} f\big(\rho\sin\theta\cos\varphi, \rho\sin\theta\sin\varphi, \rho\cos\theta\big)\,d\varphi\right)d\theta\right)d\rho.
\end{aligned} \tag{5.6.4}$$

5.7 The Divergence Theorem in $\mathbb{R}^2$

Integration by parts in more than one dimension takes many forms, and in order to even state these results in generality one needs more machinery than we have developed. Thus, we will deal with only a couple of examples and not attempt to derive a general statement.

The most basic result is a simple application of The Fundamental Theorem of Calculus, Theorem 3.2.1. Namely, consider a rectangle $R = \prod_{j=1}^{N}[a_j, b_j]$, where $N \geq 2$, and let φ be a $\mathbb{C}$-valued function which is continuously differentiable on $\text{int}(R)$ and has bounded derivatives there. Given $\boldsymbol{\xi} \in \mathbb{R}^N$, one has

$$\int_R \partial_{\boldsymbol{\xi}}\varphi(\mathbf{x})\,d\mathbf{x} = \sum_{j=1}^{N} \xi_j \left(\int_{R_j(b_j)} \varphi(\mathbf{y})\,d\sigma_{\mathbf{y}} - \int_{R_j(a_j)} \varphi(\mathbf{y})\,d\sigma_{\mathbf{y}} \right)$$
$$\text{where } R_j(c) \equiv \left(\prod_{i=1}^{j-1}[a_i, b_i] \right) \times \{c\} \times \left(\prod_{i=j+1}^{N} [a_j, b_j] \right), \tag{5.7.1}$$

where the integral $\int_{R_j(c)} \psi(\mathbf{y})\,d\sigma_{\mathbf{y}}$ of a function ψ over $R_j(c)$ is interpreted as the $(N-1)$-dimensional integral

$$\int_{\prod_{i\neq j}[a_i,b_i]} \psi(y_1, \dots, y_{j-1}, c, y_{j+1}, \dots, y_N)\,dy_1 \cdots dy_{j-1}dy_{j+1} \cdots dy_N.$$

Verification of (5.7.1) is easy. First write $\partial_{\boldsymbol{\xi}}\varphi$ as $\sum_{j=1}^{N} \xi_j \partial_{\mathbf{e}_j}\varphi$. Second, use (5.2.2) with the permutation that exchanges j and 1 but leaves the ordering of the other indices unchanged, and apply Theorem 3.2.1.

In many applications one is dealing with an $\mathbb{R}^N$-valued function $\mathbf{F}$ and is integrating its *divergence*

$$\text{div}\mathbf{F} \equiv \sum_{j=1}^{N} \partial_{\mathbf{e}_j} F_j$$

over R. By applying (5.7.1) to each coordinate, one arrives at

$$\int_R \text{div}\mathbf{F}(\mathbf{x})\,d\mathbf{x} = \sum_{j=1}^{N} \int_{R_j(b_j)} F_j(\mathbf{y})\,d\sigma_{\mathbf{y}} - \sum_{j=1}^{N} \int_{R_j(a_j)} F_j(\mathbf{y})\,d\sigma_{\mathbf{y}}, \tag{5.7.2}$$

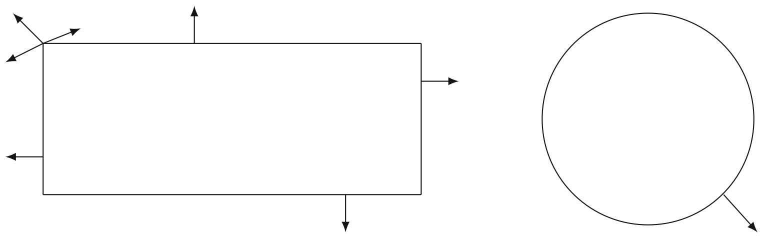
but there is a more revealing way to write (5.7.2). To explain this alternative version, let ∂G be the boundary of a bounded open subset G in $\mathbb{R}^N$. Given a point $\mathbf{x} \in \partial G$, say that $\boldsymbol{\xi} \in \mathbb{R}^N$ is a *tangent vector* to ∂G at $\mathbf{x}$ if there is a continuously differentiable path $\gamma : (-1, 1) \longrightarrow \partial G$ such that $\mathbf{x} = \gamma(0)$ and $\boldsymbol{\xi} = \dot{\gamma}(0) \equiv \frac{d\gamma}{dt}(0)$. That is, $\boldsymbol{\xi}$

is the velocity at the time when a path on ∂G passes through $\mathbf{x}$. For instance, when $G = \text{int}(R)$ and $\mathbf{x} \in R_j(a_j) \cup R_j(b_j)$ is not on one of the edges, then it is obvious that $\boldsymbol{\xi}$ is tangent to $\mathbf{x}$ if and only if $(\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N} = 0$. If $\mathbf{x}$ is at an edge, $(\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N}$ will be 0 for every tangent vector $\boldsymbol{\xi}$, but there will be $\boldsymbol{\xi}$'s for which $(\boldsymbol{\xi}, \mathbf{e}_j)_{\mathbb{R}^N} = 0$ and yet there is no continuously differentiable path that stays on ∂G, passes through $\mathbf{x}$, and has derivative $\boldsymbol{\xi}$ when it does. When $G = B(\mathbf{0}, r)$ and $\mathbf{x} \in \mathbb{S}^{N-1}(\mathbf{0}, r) \equiv \partial B(\mathbf{0}, r)$, then $\boldsymbol{\xi}$ is tangent to ∂G if and only if $(\boldsymbol{\xi}, \mathbf{x})_{\mathbb{R}^N} = 0$.[5] To see this, first suppose that $\boldsymbol{\xi}$ is tangent to ∂G at $\mathbf{x}$, and let γ be an associated path. Then

$$0 = \partial_t |\gamma(t)|^2 = 2\big(\gamma(t), \dot{\gamma}(t)\big)_{\mathbb{R}^N} = 2(\mathbf{x}, \boldsymbol{\xi})_{\mathbb{R}^N} \quad \text{at } t = 0.$$

Conversely, suppose that $(\mathbf{x}, \boldsymbol{\xi})_{\mathbb{R}^N} = 0$. If $\boldsymbol{\xi} = \mathbf{0}$, then we can take $\gamma(t) = \mathbf{x}$ for all t. If $\boldsymbol{\xi} \neq \mathbf{0}$, define $\gamma(t) = \big(\cos(r^{-1}|\boldsymbol{\xi}|t)\big)\mathbf{x} + \frac{r}{|\boldsymbol{\xi}|}\big(\sin(r^{-1}|\boldsymbol{\xi}|t)\big)\boldsymbol{\xi}$, and check that $\gamma(t) \in \mathbb{S}^{N-1}(\mathbf{0}, r)$ for all t, $\gamma(0) = \mathbf{x}$, and $\dot{\gamma}(0) = \boldsymbol{\xi}$.

Having defined what it means for a vector to be tangent to ∂G at $\mathbf{x}$, we now say that a vector $\boldsymbol{\eta}$ is a *normal vector* to ∂G at $\mathbf{x}$ if $(\boldsymbol{\eta}, \boldsymbol{\xi})_{\mathbb{R}^N} = 0$ for every tangent vector $\boldsymbol{\xi}$ at $\mathbf{x}$. For nice regions like balls, there is essentially only one normal vector at a point. Indeed, as we saw, $\boldsymbol{\xi}$ is tangent to $\mathbf{x} \in \partial B(\mathbf{0}, r)$ if and only if $(\boldsymbol{\xi}, \mathbf{x})_{\mathbb{R}^N} = 0$, and so every normal vector there will have the form $\alpha\mathbf{x}$ for some $\alpha \in \mathbb{R}$. In particular, there is a unique unit vector, known as the *outward pointing unit normal vector* $\mathbf{n}(\mathbf{x})$, that is normal to $\partial B(\mathbf{0}, r)$ at $\mathbf{x}$ and is pointing outward in the sense that $\mathbf{x} + t\mathbf{n}(\mathbf{x}) \notin B(\mathbf{0}, r)$ for $t > 0$. In fact, $\mathbf{n}(\mathbf{x}) = \frac{\mathbf{x}}{|\mathbf{x}|}$. Similarly, when $\mathbf{x} \in R_j(a_j) \cup R_j(b_j)$ is not on an edge, every normal vector will be of the form $\alpha\mathbf{e}_j$, and the outward pointing normal unit normal at $\mathbf{x}$ will be $-\mathbf{e}_j$ or $\mathbf{e}_j$ depending on whether $\mathbf{x} \in R_j(a_j)$ or $\mathbf{x} \in R_j(b_j)$. However, when $\mathbf{x}$ is at an edge, there are too few tangent vectors to uniquely determine an outward pointing unit normal vector at $\mathbf{x}$.



normal to rectangle and circle

Fortunately, because this flaw is present only on a Riemann negligible set, it is not fatal for the application that we will make of these concepts to (5.7.2). To be precise, define $\mathbf{n}(\mathbf{x}) = \mathbf{0}$ for $\mathbf{x} \in \partial R$ that are on an edge, note that $\mathbf{n}$ is continuous off of a Riemann negligible subset of $\mathbb{R}^{N-1}$, and observe that (5.7.2) can be rewritten as

[5] This fact accounts for the notation $\mathbb{S}^{N-1}$ when referring to spheres in $\mathbb{R}^N$. Such surfaces are said to be $(N-1)$-dimensional because there are only $N-1$ linearly independent directions in which one can move without leaving them.

$$\int_R \operatorname{div}\mathbf{F}(\mathbf{x})\,d\mathbf{x} = \int_{\partial R} \big(\mathbf{F}(\mathbf{y}), \mathbf{n}(\mathbf{y})\big)_{\mathbb{R}^N}\,d\sigma_{\mathbf{y}}, \tag{5.7.3}$$

where

$$\int_{\partial R} \psi(\mathbf{y})\,d\sigma_{\mathbf{y}} \equiv \sum_{j=1}^{N}\left(\int_{R_j(a_j)} \psi(\mathbf{y})\,d\sigma_{\mathbf{y}} + \int_{R_j(b_j)} \psi(\mathbf{y})\,d\sigma_{\mathbf{y}}\right).$$

Besides being more aesthetically pleasing than (5.7.2), (5.7.3) has the advantage that it is in a form that generalizes and has a nice physical interpretation. In fact, once one knows how to interpret integrals over the boundary of more general regions, one can show that

$$\int_G \operatorname{div}\big(\mathbf{F}(\mathbf{x})\big)\,d\mathbf{x} = \int_{\partial G} \big(\mathbf{F}(\mathbf{y}), \mathbf{n}(\mathbf{y})\big)_{\mathbb{R}^N}\,d\sigma_{\mathbf{y}} \tag{5.7.4}$$

holds for quite general regions, and this generalization is known as the *divergence theorem*. Unfortunately, understanding of the physical interpretation requires one to know the relationship between $\operatorname{div}\mathbf{F}$ and the *flow* that $\mathbf{F}$ determines, and, although a rigorous explanation of this connection is beyond the scope of this book, here is the idea. In Sect. 4.5 we showed that if $\mathbf{F}$ satisfies (4.5.4), then it determines a map $\mathbf{X} : \mathbb{R} \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$ by the equation

$$\dot{\mathbf{X}}(t, \mathbf{x}) = \mathbf{F}\big(\mathbf{X}(t, \mathbf{x})\big) \quad \text{with } \mathbf{X}(0, \mathbf{x}) = \mathbf{x}.$$

In other words, for each $\mathbf{x}$, $t \rightsquigarrow \mathbf{X}(t, \mathbf{x})$ is the path that passes through $\mathbf{x}$ at time $t = 0$ and has velocity $\mathbf{F}\big(\mathbf{X}(t, \mathbf{x})\big)$ for all t. Now think about mass that is initially uniformly distributed in a bounded region G and that is flowing along these paths. If one monitors the region to determine how much mass is lost or gained as a consequence of the flow, one can show that the rate at which change is taking place is given by the integral of $\operatorname{div}\mathbf{F}$ over G. If instead of monitoring the region, one monitors the boundary and measures how much mass is passing through it in each direction, then one finds that the rate of change is given by the integral of $\big(\mathbf{F}(\mathbf{x}), \mathbf{n}(\mathbf{x})\big)_{\mathbb{R}^N}$ over the boundary. Thus, (5.7.4) is simply stating that these two methods of measurement give the same answer.

We will now verify (5.7.4) for a special class of regions in $\mathbb{R}^2$. The main reason for working in $\mathbb{R}^2$ is that regions there are likely to have boundaries that are piecewise parameterized curves, which, by the results in Sect. 4.4, means that we know how to integrate over them. The regions G with which we will deal are *piecewise smooth star shaped regions in* $\mathbb{R}^2$ given by (5.6.1) with a continuous function $\varphi \in [0, 2\pi] \longmapsto r(\varphi) \in (0, \infty)$ that satisfies $r(0) = r(2\pi)$ and is piecewise smooth. Clearly the boundary of such a region is a piecewise parameterized curve. Indeed, consider the path $\mathbf{p}(\varphi) = \mathbf{c} + r(\varphi)\mathbf{e}(\varphi)$ where, as before, $\mathbf{e}(\varphi) = (\cos\varphi, \sin\varphi)$. Then the restriction $\mathbf{p}_1$ of $\mathbf{p}$ to $[0, \pi]$ and the restriction $\mathbf{p}_2$ of $\mathbf{p}$ to $[\pi, 2\pi]$ parameterize non-overlapping parameterized curves whose union is ∂G. Moreover, by (4.4.2), since

$$\dot{\mathbf{p}}(\varphi) = r'(\varphi)\mathbf{e}(\varphi) + r(\varphi)\dot{\mathbf{e}}(\varphi), \ \big(\mathbf{e}(\varphi), \dot{\mathbf{e}}(\varphi)\big)_{\mathbb{R}^2} = 0, \text{ and } |\mathbf{e}(\varphi)| = |\dot{\mathbf{e}}(\varphi)| = 1,$$

we have that

$$\int_{\partial G} f(\mathbf{y})\, d\sigma_{\mathbf{y}} = \int_0^{2\pi} f\big(\mathbf{c} + r(\varphi)\mathbf{e}(\varphi)\big)\sqrt{r(\varphi)^2 + r'(\varphi)^2}\, d\varphi.$$

Next observe that,

$$\mathbf{t}(\varphi) \equiv \big(r'(\varphi)\cos\varphi - r(\varphi)\sin\varphi, r'(\varphi)\sin\varphi + r(\varphi)\cos\varphi\big)$$

is tangent to ∂G at $\mathbf{p}(\varphi)$, and therefore that the outward pointing unit normal to ∂G at $\mathbf{p}(\varphi)$ is

$$\mathbf{n}\big(\mathbf{p}(\varphi)\big) = \pm\frac{\big(r(\varphi)\cos\varphi + r'(\varphi)\sin\varphi, r(\varphi)\sin\varphi - r'(\varphi)\cos\varphi\big)}{\sqrt{r(\varphi)^2 + r'(\varphi)^2}} \tag{5.7.5}$$

for $\varphi \in [0, 2\pi]$ in intervals where r is continuously differentiable. Further, since

$$\big(\mathbf{n}\big(\mathbf{p}(\varphi)\big), \mathbf{p}(\varphi) - \mathbf{c}\big)_{\mathbb{R}^2} = \frac{r(\varphi)^2}{\sqrt{r(\varphi)^2 + r'(\varphi)^2}} > 0$$

and therefore

$$|\mathbf{p}(\varphi) + t\mathbf{n}(\varphi) - \mathbf{c}|^2 > r(\varphi)^2 \quad \text{for } t > 0,$$

we know that the plus sign is the correct one. Taking all these into account, we see that (5.7.4) for G is equivalent to

$$\begin{aligned}\int_G \mathrm{div}\mathbf{F}(\mathbf{x})\, d\mathbf{x} = &\int_0^{2\pi} \big(r(\varphi)\cos\varphi + r'(\varphi)\sin\varphi\big) F_1\big(\mathbf{c} + r(\varphi)\mathbf{e}(\varphi)\big)\, d\varphi \\ &+ \int_0^{2\pi} \big(r(\varphi)\sin\varphi - r'(\varphi)\cos\varphi\big) F_2\big(\mathbf{c} + r(\varphi)\mathbf{e}(\varphi)\big)\, d\varphi.\end{aligned} \tag{5.7.6}$$

In proving (5.7.6), we will assume, without loss in generality, that $\mathbf{c} = \mathbf{0}$. Hence, what we have to show is that

$$(*) \qquad \begin{aligned}\int_G \partial_{\mathbf{e}_1} f(\mathbf{x})\, d\mathbf{x} &= \int_0^{2\pi} \big(r(\varphi)\cos\varphi + r'(\varphi)\sin\varphi\big) f\big(r(\varphi)\mathbf{e}(\varphi)\big)\, d\varphi \\ \int_G \partial_{\mathbf{e}_2} f(\mathbf{x})\, d\mathbf{x} &= \int_0^{2\pi} \big(r'(\varphi)\sin\varphi - r'(\varphi)\big) f\big(r(\varphi)\mathbf{e}(\varphi)\big)\, d\varphi\end{aligned}$$

To perform the required computation, it is important to write derivatives in terms of the variables ρ and φ. For this purpose, suppose that f is a continuously differentiable function on an open subset of $\mathbb{R}^2$, and set $g(\rho, \varphi) = f(\rho\cos\varphi, \rho\sin\varphi)$. Then

$$\partial_\rho g(\rho,\varphi) = \cos\varphi\,\partial_{\mathbf{e}_1} f(\rho\cos\varphi,\rho\sin\varphi) + \sin\varphi\,\partial_{\mathbf{e}_2} f(\rho\cos\varphi,\rho\sin\varphi)$$

and

$$\partial_\varphi g(\rho,\varphi) = -\rho\sin\varphi\,\partial_{\mathbf{e}_1} f(\rho\cos\varphi,\rho\sin\varphi) + \rho\cos\varphi\,\partial_{\mathbf{e}_2} f(\rho\cos\varphi,\rho\sin\varphi),$$

and therefore

$$\rho\partial_{\mathbf{e}_1} f(\rho\mathbf{e}(\varphi)) = \rho\cos\varphi\,\partial_\rho g(\rho,\varphi) - \sin\varphi\,\partial_\varphi g(\rho,\varphi)$$

and

$$\rho\partial_{\mathbf{e}_2} f(\rho\mathbf{e}(\varphi)) = \rho\sin\varphi\,\partial_\rho g(\rho,\varphi) + \cos\varphi\,\partial_\varphi g(\rho,\varphi).$$

Thus, if f is a continuous function on $\bar{G}$ that has bounded, continuous first order derivatives on G, then

$$\int_G \partial_{\mathbf{e}_1} f(\mathbf{x})\,d\mathbf{x} = I - J,$$

where

$$I = \int_0^{2\pi} \cos\varphi \left(\int_0^{r(\varphi)} \rho\partial_\rho g(\rho,\theta)\,d\rho\right) d\varphi$$

and

$$J = \int_0^{2\pi} \sin\varphi \left(\int_0^{r(\varphi)} \partial_\varphi g(\rho,\theta)\,d\rho\right) d\varphi.$$

Applying integration by parts to the inner integral in I, we see that

$$I = \int_0^{2\pi} \cos\varphi\, g\big(r(\varphi),\varphi\big) r(\varphi)\,d\varphi - \int_0^{2\pi} \cos\varphi \left(\int_0^{r(\varphi)} g(\rho,\varphi)\,d\rho\right) d\varphi.$$

Dealing with J is more challenging. The first step is to write it as $J_1 + J_2$, where J_1 and J_2 are, respectively,

$$\int_0^{2\pi} \sin\varphi \left(\int_0^{r_-} \partial_\varphi g(\rho,\varphi)\,d\rho\right) d\varphi \text{ and } \int_0^{2\pi} \sin\varphi \left(\int_{r_-}^{r(\varphi)} \partial_\varphi g(\rho,\varphi)\,d\rho\right) d\varphi,$$

and $r_- \equiv \min\{r(\varphi) : \varphi \in [0,2\pi]\}$. By (5.2.1) and integration by parts,

$$J_1 = \int_0^{r_-} \left(\int_0^{2\pi} \sin\varphi\,\partial_\varphi g(\rho,\varphi)\,d\varphi\right) d\rho = -\int_0^{r_-} \left(\int_0^{2\pi} \cos\varphi\, g(\rho,\varphi)\,d\varphi\right) d\rho.$$

To handle J_2, choose $\theta_0 \in [0, 2\pi]$ so that $r(\theta_0) = r_-$, and choose $\theta_1, \dots, \theta_\ell \in [0, 2\pi]$ so that r' is continuous on each of the open intervals with end points θ_k and θ_{k+1}, where $\theta_{\ell+1} = \theta_0$. Now use (3.1.6) to write J_2 as $\sum_{k=0}^{\ell} J_{2,k}$, where

$$J_{2,k} = \int_0^{2\pi} \sin\varphi \left(\int_{r(\varphi\wedge\theta_k)}^{r(\varphi\wedge\theta_{k+1})} \partial_\varphi g(\rho, \varphi)\, d\rho \right) d\varphi,$$

and then make the change of variables $\rho = r(\theta)$ and apply (5.2.1) to obtain

$$\begin{aligned} J_{2,k} &= \int_0^{2\pi} \sin\varphi \left(\int_{\varphi\wedge\theta_k}^{\varphi\wedge\theta_{k+1}} \partial_\varphi g\big(r(\theta), \varphi\big) r'(\theta)\, d\theta \right) d\varphi \\ &= \int_{\theta_k}^{\theta_{k+1}} r'(\theta) \left(\int_\theta^{2\pi} \sin\varphi\, \partial_\varphi g\big(r(\theta), \varphi\big)\, d\varphi \right) d\theta. \end{aligned}$$

Hence,

$$J_2 = \int_0^{2\pi} r'(\theta) \left(\int_\theta^{2\pi} \sin\varphi\, \partial_\varphi g\big(r(\theta), \varphi\big)\, d\varphi \right) d\theta,$$

which, after integration by parts is applied to the inner integral, leads to

$$J_2 = -\int_0^{2\pi} \sin\theta\, g\big(r(\theta), \theta\big) r'(\theta)\, d\theta - \int_0^{2\pi} r'(\theta) \left(\int_\theta^{2\pi} \cos\varphi\, g\big(r(\theta), \varphi\big)\, d\varphi \right) d\theta.$$

After applying (5.2.1) and undoing the change of variables in the second integral on the right, we get

$$J_2 = -\int_0^{2\pi} \sin\theta\, g\big(r(\theta), \theta\big) r'(\theta)\, d\theta - \int_0^{2\pi} \cos\varphi \left(\int_{r_-}^{r(\varphi)} g(\rho, \varphi)\, d\rho \right) d\varphi$$

and therefore

$$J = -\int_0^{2\pi} \sin\varphi\, g\big(r(\varphi), \varphi\big) r'(\varphi)\, d\varphi - \int_0^{2\pi} \cos\varphi \left(\int_0^{r(\varphi)} g(\rho, \varphi)\, d\rho \right) d\varphi.$$

Finally, when we subtract J from I, we arrive at

$$\int_G \partial_{\mathbf{e}_1} f(\mathbf{x})\, d\mathbf{x} = \int_0^{2\pi} \big(r(\varphi)\cos\varphi + r'(\varphi)\sin\varphi\big) g\big(r(\varphi), \varphi\big)\, d\varphi.$$

Proceeding in exactly the same way, one can derive the second equation in $(*)$, and so we have proved the following theorem.

Theorem 5.7.1 *If $G \subseteq \mathbb{R}^2$ is a piecewise smooth star shaped region and if $\mathbf{F} : \bar{G} \longrightarrow \mathbb{R}^2$ is continuous on $\bar{G}$ and has bounded, continuous first order derivatives on G, then (5.7.5), and therefore (5.7.4) with $\mathbf{n} : \partial G \longrightarrow \mathbb{S}^1(0,1)$ given by (5.7.5), hold.*

Corollary 5.7.2 *Let G be as in Theorem 5.7.1, and suppose that $\mathbf{a}_1, \dots, \mathbf{a}_\ell \in G$ and $r_1, \dots, r_\ell \in (0,\infty)$ have the properties that $\overline{B(\mathbf{a}_k, r_k)} \subseteq G$ for each $1 \le k \le \ell$ and that $\overline{B(\mathbf{a}_k, r_k)} \cap \overline{B(\mathbf{a}_{k'}, r_{k'})} = \emptyset$ for $1 \le k < k' \le \ell$. Set $H = G \setminus \bigcup_{k=1}^{\ell} \overline{B(\mathrm{a}_k, r_k)}$. If $\mathbf{F} : \bar{H} \longrightarrow \mathbb{R}^2$ is a continuous function that has bounded, continuous first order derivatives on H, then $\int_H \mathrm{div}\mathbf{F}(\mathbf{x})\, d\mathbf{x}$ equals*

$$\int_{\partial G} \big(\mathbf{F}(\mathbf{y}), \mathbf{n}(\mathbf{y})\big)_{\mathbb{R}^2}\, d\sigma_{\mathbf{y}} \\ - \sum_{k=1}^{\ell} r_k \int_0^{2\pi} \Big(F_1\big(\mathbf{a}_k + r_k \mathbf{e}(\varphi)\big) \cos\varphi + F_2\big(\mathbf{a}_k + r_k \mathbf{e}(\varphi)\big) \sin\varphi \Big)\, d\varphi.$$

Proof First assume that $\mathbf{F}$ has bounded, continuous derivatives on the whole of G. Then Theorem 5.7.1 applies to $\mathbf{F}$ on $\bar{G}$ and its restriction to each ball $\overline{B(\mathbf{a}_k, r_k)}$, and so the result follows from Theorem 5.7.1 when one writes the integral of $\mathrm{div}\mathbf{F}$ over $\bar{H}$ as

$$\int_{\bar{G}} \mathrm{div}\mathbf{F}(\mathbf{x})\, d\mathbf{x} - \sum_{k=1}^{\ell} \int_{\overline{B(\mathbf{a}_k, r_k)}} \mathrm{div}\mathbf{F}(\mathbf{x})\, d\mathbf{x}.$$

To handle the general case, define $\eta : \mathbb{R} \longrightarrow [0,1]$ by

$$\eta(t) = \begin{cases} 0 & \text{if } t \le 0 \\ \frac{1+\sin\left(\pi(t-\frac{1}{2})\right)}{2} & \text{if } 0 < t \le 1 \\ 1 & \text{if } t > 1. \end{cases}$$

Then η is continuously differentiable. For each $1 \le k \le \ell$, choose $R_k > r_k$ so that $\overline{B(\mathbf{a}_k, R_k)} \subseteq G$ and $\overline{B(\mathbf{a}_k, R_k)} \cap \overline{B(\mathbf{a}_{k'}, R_{k'})} = \emptyset$ for $1 \le k < k' \le \ell$. Define

$$\psi_k(\mathbf{x}) = \eta\left(\frac{|\mathbf{x} - \mathbf{a}_k|^2 - r_k^2}{R_k^2 - r_k^2} \right) \quad \text{for } \mathbf{x} \in \mathbb{R}^2$$

and

$$\tilde{\mathbf{F}}(\mathbf{x}) = \sum_{k=1}^{\ell} \psi_k(\mathbf{x}) \mathbf{F}(\mathbf{x})$$

if $\mathbf{x} \in \bar{H}$ and $\tilde{\mathbf{F}}(\mathbf{x}) = 0$ if $\mathbf{x} \in \bigcup_{k=1}^{\ell} B(\mathbf{a}_k, r_k)$. Then $\tilde{\mathbf{F}}$ is continuous on $\bar{G}$ and has bounded, continuous first order derivatives on G. In addition, $\tilde{\mathbf{F}} = \mathbf{F}$ on $G \setminus \bigcup_{k=1}^{\ell} \overline{B(\mathbf{a}_k, R_k)}$. Hence, if $H' = \bar{G} \setminus \bigcup_{k=1}^{\ell} \overline{B(\mathbf{a}_k, R_k)}$, then, by the preceding, $\int_{H'} \operatorname{div}\mathbf{F}(\mathbf{x})\, d\mathbf{x}$ equals

$$\int_{\partial G} \big(\mathbf{F}(\mathbf{y}), \mathbf{n}(\mathbf{y})\big)_{R^2} \\ - \sum_{k=1}^{\ell} R_k \int_0^{2\pi} \Big(F_1\big(\mathbf{a}_k + R_k \mathbf{e}(\varphi)\big) \cos\varphi + F_2\big(\mathbf{a}_k + R_k \mathbf{e}(\varphi)\big) \sin\varphi\Big)\, d\varphi,$$

and so the asserted result follows after one lets each R_k decrease to r_k. □

We conclude this section with an application of Theorem 5.7.1 that plays a role in many places. One of the consequence of the Fundamental Theorem of Calculus is that every continuous function f on an interval (a, b) is the derivative of a continuously differentiable function F on (a, b). Indeed, simply set $c = \frac{a+b}{2}$ and take $F(x) = \int_c^x f(t)\, dt$. With this in mind, one should ask whether an analogous statement holds in $\mathbb{R}^2$. In particular, given a connected open set $G \subseteq \mathbb{R}^2$ and a continuous function $\mathbf{F} : G \longrightarrow \mathbb{R}^2$, is it true that there is a continuously differentiable function $f : G \longrightarrow \mathbb{R}$ such that $\mathbf{F}$ is the gradient of f? That the answer is *no* in general can be seen by assuming that $\mathbf{F}$ is continuously differentiable and noticing that if f exists then

$$\partial_{\mathbf{e}_2} F_1 = \partial_{\mathbf{e}_2} \partial_{\mathbf{e}_1} f = \partial_{\mathbf{e}_1} \partial_{\mathbf{e}_2} f = \partial_{\mathbf{e}_1} F_2.$$

Hence, a necessary condition for the existence of f is that $\partial_{\mathbf{e}_2} F_1 = \partial_{\mathbf{e}_1} F_2$, and when this condition holds $\mathbf{F}$ is said to be *exact*. It is known that exactness is sufficient as well as necessary for a continuously differentiable $\mathbf{F}$ on G to be the gradient of a function when G is what is called a simply connected region, but to avoid technical difficulties, we will restrict ourselves to star shaped regions.

Corollary 5.7.3 *Assume that G is a star shaped region in $\mathbb{R}^2$ and that $\mathbf{F} : G \longrightarrow \mathbb{R}^2$ is a continuously differentiable function. Then there exists a continuously differentiable function $f : G \longrightarrow \mathbb{R}$ such that $\mathbf{F} = \nabla f$ if and only if $\mathbf{F}$ is exact.*

Proof Without loss in generality, we will assume that $\mathbf{0}$ is the center of G. Further, since the necessity has already been shown, we will assume that $\mathbf{F}$ is exact.

Define $f : G \longrightarrow \mathbb{R}$ by $f(\mathbf{0}) = 0$ and

$$f\big(r\mathbf{e}(\varphi)\big) = \int_0^r \big(F_1(\rho\mathbf{e}(\varphi)) \cos\varphi + F_2(\rho\mathbf{e}(\varphi)) \sin\varphi\big)\, d\rho$$

for $\varphi \in [0, 2\pi)$ and $0 < r < r(\varphi)$. Clearly $F_1(\mathbf{0}) = \partial_{\mathbf{e}_1} f(\mathbf{0})$ and $F_2(\mathbf{0}) = \partial_{\mathbf{e}_2} f(\mathbf{0})$.

We will now show that $F_1 = \partial_{\mathbf{e}_1} f$ at any point $(\xi_0, \eta_0) \in G \setminus \{\mathbf{0}\}$. This is easy when $\eta_0 = 0$, since $f(\xi, 0) = \int_0^\xi F_1(t, 0)\, dt$. Thus assume that $\eta_0 \neq 0$, and consider points

(ξ, η_0) not equal to (ξ_0, η_0) but sufficiently close that $(t, \eta_0) \in G$ if $\xi\wedge\xi_0 \le t \le \xi\vee\xi_0$. What we need to show is that

$$(*) \qquad f(\xi, \eta_0) - f(\xi_0, \eta_0) = \int_{\xi_0}^{\xi} F_1(t, \eta_0)\, dt.$$

To this end, define $\tilde{\mathbf{F}} = (F_2, -F_1)$. Then, because $\mathbf{F}$ is exact, $\operatorname{div}\tilde{\mathbf{F}} = 0$ on G. Next consider the region H that is the interior of the triangle whose vertices are $\mathbf{0}$, (ξ_0, η_0), and (ξ, η_0). Then H is a piecewise smooth star shaped region and so, by Theorem 5.7.1, the integral of $(\tilde{\mathbf{F}}, \mathbf{n})_{\mathbb{R}^2}$ over ∂H is 0. Thus, if we write (ξ_0, η_0) and (ξ, η_0) as $r_0\mathbf{e}(\varphi_0)$ and $r\mathbf{e}(\varphi)$, then

$$\begin{aligned} 0 = \int_{\partial H} \big(\tilde{\mathbf{F}}(\mathbf{y}), \mathbf{n}(\mathbf{y})\big)_{\mathbb{R}^2}\, d\mathbf{y} = \int_0^{r_0} &\Big(\tilde{\mathbf{F}}\big(\rho\mathbf{e}(\varphi_0)\big), \mathbf{n}\big(\rho\mathbf{e}(\varphi_0)\big)\Big)_{\mathbb{R}^2} d\rho \\ &+ \int_0^{r} \Big(\tilde{\mathbf{F}}\big(\rho\mathbf{e}(\varphi)\big), \mathbf{n}\big(\rho\mathbf{e}(\varphi)\big)\Big)_{\mathbb{R}^2} d\rho \\ &+ \int_{\xi\wedge\xi_0}^{\xi\vee\xi_0} \big(\tilde{\mathbf{F}}(t, \eta_0), \mathbf{n}(t, \eta_0)\big)_{\mathbb{R}^2}\, dt. \end{aligned}$$

If $\eta_0 > 0$ and $\xi > \xi_0$, then

$$\mathbf{n}\big(\rho\mathbf{e}(\varphi)\big) = (\sin\varphi, -\cos\varphi), \;\; \mathbf{n}\big(\rho\mathbf{e}(\varphi_0)\big) = (-\sin\varphi_0, \cos\varphi_0),$$

and $\mathbf{n}(t, \eta_0) = (0, 1)$, and therefore

$$\begin{aligned} f(\xi, \eta_0) &= \int_0^{r} \Big(\tilde{\mathbf{F}}\big(\rho\mathbf{e}(\varphi)\big), \mathbf{n}\big(\rho\mathbf{e}(\varphi)\big)\Big)_{\mathbb{R}^2} d\rho, \\ f(\xi_0, \eta_0) &= -\int_0^{r_0} \Big(\tilde{\mathbf{F}}\big(\rho\mathbf{e}(\varphi_0)\big), \mathbf{n}\big(\rho\mathbf{e}(\varphi_0)\big)\Big)_{\mathbb{R}^2} d\rho, \\ \int_{\xi_0}^{\xi} F_1(t, \eta_0)\, dt &= -\int_{\xi\wedge\xi_0}^{\xi\vee\xi_0} \big(\tilde{\mathbf{F}}(t, \eta_0), \mathbf{n}(t, \eta_0)\big)_{\mathbb{R}^2}\, dt, \end{aligned}$$

and so $(*)$ holds. If $\eta_0 > 0$ and $\xi < \xi_0$, then the sign of $\mathbf{n}$ changes in each term, and therefore we again get $(*)$, and the cases when $\eta_0 < 0$ are handled similarly.

The proof that $F_2 = \partial_{\mathbf{e}_2} f$ follows the same line of reasoning and is left as an exercise. □

5.8 Exercises

Exercise 5.1 Let $\mathbf{x}, \ \mathbf{y} \in \mathbb{R}^2\backslash\{\mathbf{0}\}$. Then Schwarz's inequality says that the ratio

$$\rho \equiv \frac{(\mathbf{x}, \mathbf{y})_{\mathbb{R}^2}}{|\mathbf{x}||\mathbf{y}|}$$

is in the open interval $(-1, 1)$ unless $\mathbf{x}$ and $\mathbf{y}$ lie on the same line, in which case $\rho \in \{-1, 1\}$. Euclidean geometry provides a good explanation for this. Indeed, consider the triangle whose vertices are $\mathbf{0}$, $\mathbf{x}$, and $\mathbf{y}$. The sides of this triangle have lengths $|\mathbf{x}|$, $|\mathbf{y}|$, and $|\mathbf{y}-\mathbf{x}|$. Thus, by the law of the cosine, $|\mathbf{y}-\mathbf{x}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 - 2|\mathbf{x}||\mathbf{y}|\cos\theta$, where θ is the angle in the triangle between $\mathbf{x}$ and $\mathbf{y}$. Use this to show that $\rho = \cos\theta$. The same explanation applies in higher dimensions since there is a plane in which $\mathbf{x}$ and $\mathbf{y}$ lie and the analysis can be carried out in that plane.

Exercise 5.2 Show that

$$\int_{\mathbb{R}} e^{\lambda x} e^{-\frac{x^2}{2}}\, dx = \sqrt{2\pi} e^{\frac{\lambda^2}{2}} \quad \text{for all } \lambda \in \mathbb{R}.$$

One way to do this is to make the change of variables $y = x - \lambda$ to see that

$$\int_{\mathbb{R}} e^{\lambda x} e^{-\frac{x^2}{2}}\, dx = e^{\frac{\lambda^2}{2}} \int_{\mathbb{R}} e^{-\frac{(y-\lambda)^2}{2}}\, dy,$$

and then use (5.1.2) and (5.4.3).

Exercise 5.3 Show that $|\Gamma|_e = |\mathfrak{R}(\Gamma)|_e$ and $|\Gamma|_i = |\mathfrak{R}(\Gamma)|_i$ for all bounded sets $\Gamma \subseteq \mathbb{R}^N$ and all rotations $\mathfrak{R}$, and conclude that $\mathfrak{R}(\Gamma)$ is Riemann measurable if and only if Γ is. Use this to show that if Γ is a bounded subset of $\mathbb{R}^N$ for which there exists a $\mathbf{x}_0 \in \mathbb{R}^N$ and an $\mathbf{e} \in \mathbb{S}^{N-1}(\mathbf{0}, 1)$ such that $(\mathbf{x} - \mathbf{x}_0, \mathbf{e})_{\mathbb{R}^N} = 0$ for all $\mathbf{x} \in \Gamma$, then Γ is Riemann negligible.

Exercise 5.4 An integral that arises quite often is one of the form

$$I(a, b) \equiv \int_{(0,\infty)} t^{-\frac{1}{2}} e^{-a^2 t - \frac{b^2}{t}}\, dt,$$

where $a, b \in (0, \infty)$. To evaluate this integral, make the change of variables $\xi = at^{\frac{1}{2}} - bt^{-\frac{1}{2}}$. Then $\xi^2 = at - 2ab + t^{-1}$ and $t^{\frac{1}{2}} = \frac{\xi + \sqrt{\xi^2 + 4ab}}{2a}$, the plus being dictated by the requirement that $t \ge 0$. After making this substitution, arrive at

$$I(a, b) = \frac{e^{-2ab}}{a} \int_{\mathbb{R}} e^{-\xi^2} \left(1 + (\xi^2 + 4ab)^{-\frac{1}{2}} \xi\right) d\xi = \frac{e^{-2ab}}{a} \int_{\mathbb{R}} e^{-\xi^2}\, d\xi,$$

from which it follows that $I(a, b) = \frac{\pi^{\frac{1}{2}} e^{-2ab}}{a}$. Finally, use this to show that

$$\int_{(0,\infty)} t^{-\frac{3}{2}} e^{-a^2 t - \frac{b^2}{t}}\, dt = \frac{\pi^{\frac{1}{2}} e^{-2ab}}{b}.$$

Exercise 5.5 Recall the cardiod described in Exercise 2.2, and consider the region that it encloses. Using the expression $z(\theta) = 2R(1-\cos\theta)e^{i\theta}$ for the boundary of this region after translating by $-R$, show that the area enclosed by the cardiod is $6\pi R^2$. Finally, show that the arc length of the cardiod is $16R$, a computation in which you may want to use (1.5.1) to write $1-\cos\theta$ as a square.

Exercise 5.6 Given $a_1,\ a_2,\ a_3 \in (0,\infty)$, consider the region Ω in $\mathbb{R}^3$ enclosed by the ellipsoid $\sum_{i=1}^3 \frac{x_i^2}{a_i^2} = 1$. Show that the boundary of Ω is Riemann negligible and that the volume of Ω is $\frac{4\pi a_1 a_2 a_3}{3}$. When computing the volume V, first show that

$$V = 2a_3 \int_{\tilde{\Omega}} \sqrt{1 - \frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2}}\, dx_1 dx_2 \quad \text{where } \tilde{\Omega} = \left\{\mathbf{x} \in \mathbb{R}^2 : \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} \le 1\right\}.$$

Next, use Fubini's Theorem and a change of variables in each of the coordinates to write the integral as $a_1 a_2 \int_{B(\mathbf{0},1)} \sqrt{1-|\mathbf{x}|^2}\, d\mathbf{x}$, where $B(\mathbf{0},1)$ is the unit ball in $\mathbb{R}^2$. Finally, use (5.6.2) to complete the computation.

Exercise 5.7 Let Ω be a bounded, closed region in $\mathbb{R}^3$ with Riemann negligible boundary, and let $\mu : \Omega \longrightarrow [0,\infty)$ be a continuous function. Thinking of μ as a mass density, one says that the *center of gravity* of Ω with mass distribution μ is the point $\mathbf{c} \in \mathbb{R}^3$ such that $\int_\Omega \mu(\mathbf{y})(\mathbf{y}-\mathbf{c})\, d\mathbf{y} = \mathbf{0}$. The reason for the name is that if Ω is supported at this point $\mathbf{c}$, then the net effect of gravity will be 0 and so the region will be balanced there. Of course, $\mathbf{c}$ need not lie in Ω, in which case one should think of Ω being attached to $\mathbf{c}$ by weightless wires.

Obviously, $\mathbf{c} = \frac{\int_\Omega \mu(\mathbf{y})\mathbf{y}\, d\mathbf{y}}{M}$, where $M = \int_\Omega \mu(\mathbf{y})\, d\mathbf{y}$ is the total mass. Now suppose that $\Omega = \{\mathbf{y} \in \mathbb{R}^3 : x_1^2 + x_2^2 \le x_3 \le h\}$, where $h > 0$, has a constant mass density. Show that $\mathbf{c} = \left(0, 0, \frac{3h}{4}\right)$.

Exercise 5.8 We showed that for a ball $\overline{B(\mathbf{c},r)}$ in $\mathbb{R}^3$ with a continuous mass distribution that depends only of the distance to $\mathbf{c}$, the gravitational force it exerts on a particle of mass m at a point $\mathbf{b} \notin \overline{B(\mathbf{c},r)}$ is given by (5.6.3). Now suppose that $\mathbf{b} \in \overline{B(\mathbf{c},r)}$, and set $D = |\mathbf{b}-\mathbf{c}|$. Show that

$$\int_{\overline{B(\mathbf{c},r)}} \frac{Gm\mu(|\mathbf{y}-\mathbf{c}|)}{|\mathbf{y}-\mathbf{b}|}\, d\mathbf{y} = \frac{GmM_D}{D^2}(\mathbf{c}-\mathbf{b})$$
$$\text{where } M_D = \int_{\overline{B(\mathbf{c},D)}} \mu(|\mathbf{y}-\mathbf{c}|)\, d\mathbf{y}.$$

In other words, the forces produced by the mass that lies further than $\mathbf{b}$ from $\mathbf{c}$ cancel out, and so the particle feels only the force coming from the mass between it and the center of the ball.

Exercise 5.9 Let $B(\mathbf{0}, r)$ be the ball of radius r in $\mathbb{R}^N$ centered at the origin. Using rotation invariance, show that

$$\int_{B(\mathbf{0},r)} x_i \, d\mathbf{x} = 0 \quad \text{and} \quad \int_{B(\mathbf{0},r)} x_i x_j \, d\mathbf{x} = \frac{\Omega_N r^{N+2}}{N+2} \delta_{i,j} \text{ for } 1 \le i, j \le N.$$

Next, suppose that $f : \mathbb{R}^N \longrightarrow \mathbb{R}$ is twice continuously differentiable, and let $\mathcal{A}(f, r) = (\Omega_N r^N)^{-1} \int_{B(\mathbf{0},r)} f(\mathbf{x}) \, d\mathbf{x}$ be the average value of f on $B(\mathbf{0}, r)$. As an application of the preceding, show that $\frac{A(f,r) - f(\mathbf{0})}{r^2} \longrightarrow \frac{1}{2(N+2)} \sum_{i=1}^N \partial_{\mathbf{e}_i}^2 f(\mathbf{0})$ as $r \searrow 0$.

Exercise 5.10 Suppose that G is an open subset of $\mathbb{R}^2$ and that $\mathbf{F} : G \longrightarrow \mathbb{R}^2$ is continuous. If $\mathbf{F} = \nabla f$ for some continuously differentiable $f : G \longrightarrow \mathbb{R}$ and $\mathbf{p} : [a, b] \longrightarrow G$ is a piecewise smooth path, show that

$$f(b) - f(a) = \int_a^b \Big(\mathbf{F}\big(\mathbf{p}(t)\big), \dot{\mathbf{p}}(t)\Big)_{\mathbb{R}^N} dt$$

and therefore that $\int_a^b \Big(\mathbf{F}\big(\mathbf{p}(t)\big), \dot{\mathbf{p}}(t)\Big)_{\mathbb{R}^N} dt = 0$ if $\mathbf{p}$ is closed (i.e., $\mathbf{p}(b) = \mathbf{p}(a)$). Now assume that G is connected and that $\int_a^b \Big(\mathbf{F}\big(\mathbf{p}(t)\big), \dot{\mathbf{p}}(t)\Big)_{\mathbb{R}^N} dt = 0$ for all piecewise smooth, closed paths $\mathbf{p} : [a, b] \longrightarrow G$. Using Exercise 4.5, show that for each $\mathbf{x}, \mathbf{y} \in G$ there is a piecewise smooth path in G that starts at $\mathbf{x}$ and ends at $\mathbf{y}$. Given a reference point $\mathbf{x}_0 \in G$ and an $\mathbf{x} \in G$, show that

$$f(\mathbf{x}) \equiv \int_a^b \Big(\mathbf{F}\big(\mathbf{p}(t)\big), \dot{\mathbf{p}}(t)\Big)_{\mathbb{R}^N} dt$$

is the same for all piecewise smooth paths $\mathbf{p} : [a, b] \longrightarrow G$ such that $\mathbf{p}(a) = \mathbf{x}_0$ and $\mathbf{p}(b) = \mathbf{x}$. Finally, show that f is continuously differentiable and that $\mathbf{F} = \nabla f$.

Chapter 6
A Little Bit of Analytic Function Theory

In this concluding chapter we will return to the topic introduced in Chap. 2, and, as we will see, the divergence theorem will allow us to prove results here that were out of reach earlier.

Before getting started, now that we have introduced directional derivatives, it will be useful to update the notation used in Chap. 2. In particular, what we denoted by f'_0 and $f'_{\frac{\pi}{2}}$ there are the directional derivatives of f in the directions $\mathbf{e}_1$ and $\mathbf{e}_2$. However, because we will be identifying $\mathbb{R}^2$ with $\mathbb{C}$ and will be writing $z \in \mathbb{C}$ as $z = x + iy$, in the following we will use $\partial_x f$ and $\partial_y f$ to denote $f'_0 = \partial_{\mathbf{e}_1} f$ and $f'_{\frac{\pi}{2}} = \partial_{\mathbf{e}_2} f$. Thus, for example, in this notation, the Cauchy–Riemann equations in (2.3.1) become

$$\partial_x u = \partial_y v \quad \text{and} \quad \partial_x v = -\partial_y u. \tag{6.0.1}$$

In this connection, it will be convenient to introduce the operator given by $\bar{\partial} = \frac{1}{2}(\partial_x + i\partial_y)$ (i.e., $2\bar{\partial} f = \partial_x f + i\partial_y f$) and to notice that if $f = u + iv$ then $2\bar{\partial} f = (\partial_x u - \partial_y v) + i\big(\partial_x v + \partial_y u\big)$. Hence u and v satisfy (6.0.1) if and only if $\bar{\partial} f = 0$. Equivalently, f is analytic if and only if $\bar{\partial} f = 0$. As a consequence, it is easy to check the properties of analytic functions discussed in Exercise 2.7.

6.1 The Cauchy Integral Formula

This section is devoted to the derivation of a formula, discovered by Cauchy, which is the key to everything that follows. There are several direct proofs of his formula, but we will give one that is based on (5.7.4). For this purpose, suppose that $f = u + iv$, and observe that $2\bar{\partial} f = \operatorname{div}\mathbf{F} + i \operatorname{div}\tilde{\mathbf{F}}$, where $\mathbf{F} = \big(u, -v\big)$ and $\tilde{\mathbf{F}} = \big(v, u\big)$. Hence, if G is the piecewise smooth star shaped region in (5.6.1), and $f : \bar{G} \longrightarrow \mathbb{C}$ is a continuous function that has a bounded, continuous first order derivatives on G, then, by Theorem 5.7.1,

D.W. Stroock, *A Concise Introduction to Analysis*,
DOI 10.1007/978-3-319-24469-3_6

$$2\int_{\bar{G}} \bar{\partial} f(\mathbf{x})\, d\mathbf{x} = \int_{\partial G} \Big(\mathbf{F}(\mathbf{y}) + i\tilde{\mathbf{F}}(\mathbf{y}), \mathbf{n}(\mathbf{y})\Big)_{\mathbb{R}^2} d\sigma_{\mathbf{y}}.$$

Now let $\mathbf{c}$ be the center of G and $r : [0, 2\pi] \longrightarrow (0, \infty)$ the associated radial function (cf. (5.6.1)). Then, by using (5.7.5), we can write the integral on the right hand side of the preceding as

$$\int_0^{2\pi} \bigg(\Big(u\big(\mathbf{c} + r(\varphi)\mathbf{e}(\varphi)\big) + iv\big(\mathbf{c} + r(\varphi)\mathbf{e}(\varphi)\big)\Big)\big(r(\varphi)\cos\varphi + r'(\varphi)\sin\varphi\big) \\ + \Big(-v\big(\mathbf{c} + r(\varphi)\mathbf{e}(\varphi)\big) + iu\big(\mathbf{c} + r(\varphi)\mathbf{e}(\varphi)\big)\Big)\big(r(\varphi)\sin\varphi - r'(\varphi)\cos\varphi\big)\bigg) d\varphi,$$

where $\mathbf{e}(\varphi) \equiv (\cos\varphi, \sin\varphi)$. Thinking of $\mathbb{R}^2$ as $\mathbb{C}$ and taking $c = c_1 + ic_2$, the preceding becomes

$$\int_0^{2\pi} f\big(c + r(\varphi)e^{i\varphi}\big)\Big(\big(r(\varphi)\cos\varphi + r'(\varphi)\sin\varphi\big) + i\big(r(\varphi)\sin\varphi - r'(\varphi)\cos\varphi\big)\Big) d\varphi.$$

Finally, observing that

$$\frac{d}{d\varphi} r(\varphi)e^{i\varphi} = i\Big(\big(r(\varphi)\cos\varphi + r'(\varphi)\sin\varphi\big) + i\big(r(\varphi)\sin\varphi - r'(\varphi)\cos\varphi\big)\Big),$$

we arrive at

$$2i\int_G \bar{\partial} f(z)\, dx\, dy = \int_0^{2\pi} f\big(z_G(\varphi)\big) z_G'(\varphi)\, d\varphi, \tag{6.1.1}$$

where $z_G(\varphi) = c + r(\varphi)e^{i\varphi}$ and, as we will throughout, we have used $dx\,dy$ instead of $d\mathbf{x} = dx_1 dx_2$ to indicate Riemann integration on $\mathbb{R}^2$ when we are thinking of $\mathbb{R}^2$ as $\mathbb{C}$.

Suppose that G and $\{D(a_k, r_k) : 1 \le k \le \ell\}$ (we are thinking of $\mathbb{R}^2$ as $\mathbb{C}$, and so the ball $B(\mathbf{a}_k, r_k)$ in $\mathbb{R}^2$ is replaced by the disk $D(a_k, r_k)$ in $\mathbb{C}$ where a_k is the complex number determined by $\mathbf{a}_k$) are as in Corollary 5.7.2. The same line of reasoning that led to (6.1.1) shows that if $H = G\setminus \bigcup_{k=1}^{\ell} \overline{D(a_k, r_k)}$ then, again with $z_G(\varphi) = c + r(\varphi)e^{i\varphi}$,

$$2i\int_H \bar{\partial} f(z)\, dx\, dy = \int_0^{2\pi} f\big(z_G(\varphi)\big) z_G'(\varphi)\, d\varphi \\ - i\sum_{k=1}^{\ell} r_k \int_0^{2\pi} f\big(a_k + r_k e^{i\varphi}\big) e^{i\varphi}\, d\varphi \tag{6.1.2}$$

for continuous functions $f : \bar{H} \longrightarrow \mathbb{C}$ that have bounded, continuous first order derivatives on H.

Theorem 6.1.1 *Let G be the piecewise smooth star shaped region given by (5.6.1) with center c and radial function $r : [0, 2\pi] \longrightarrow (0, \infty)$, and define $z_G(\varphi) = c + r(\varphi)e^{i\varphi}$ as in (6.1.1). If $f : \bar{G} \longrightarrow \mathbb{C}$ is a continuous function which is analytic on G, then*

$$\int_0^{2\pi} f\big(z_G(\varphi)\big)z'_G(\varphi)\,d\varphi = 0. \tag{6.1.3}$$

Furthermore, if $z_1, \dots, z_\ell \in G$ and $r_1, \dots, r_\ell > 0$ satisfy $\overline{D(z_k, r_k)} \subseteq G$ and $\overline{D(z_k, r_k)} \cap \overline{D(z_{k'}, r_{k'})} = \emptyset$ for all $1 \le k < k' \le \ell$, then for any continuous $f : \bar{G} \setminus \bigcup_{k=1}^{\ell} D(z_k, r_k) \longrightarrow \mathbb{C}$ which is analytic on $G \setminus \bigcup_{k=1}^{\ell} \overline{D(z_k, r_k)}$,

$$\int_a^b f\big(z_G(\varphi)\big)z'_G(\varphi)\,d\varphi = i\sum_{k=1}^{\ell} r_k \int_0^{2\pi} f\big(z_k + r_k e^{i\varphi}\big)e^{i\varphi}\,d\varphi. \tag{6.1.4}$$

Proof Clearly (6.1.3) and (6.1.4) follow from (6.1.1) and (6.1.2), respectively, when the first derivatives of f have continuous extensions to $\bar{G} \setminus \bigcup_{k=1}^{\ell} D(z_k, r_k)$. More generally, choose $0 < \delta < \min\{r(\varphi) : \varphi \in [0, 2\pi]\}$ so that $\overline{D(z_k, r_k + \delta)} \cap \overline{D(z_{k'}, r_{k'} + \delta)} = \emptyset$ for $1 \le k < k' \le \ell$ and

$$\bigcup_{k=1}^{\ell} \overline{D(z_k, r_k + \delta)} \subseteq G_\delta \equiv \{c + re^{i\varphi} : \varphi \in [0, 2\pi] \;\&\; 0 \le r < r(\varphi) - \delta\}.$$

Then (6.1.3) and (6.1.4) hold with G replaced by G_δ and each $D(z_k, r_k)$ replaced by $D(z_k, r_k + \delta)$. Finally, let $\delta \searrow 0$. □

The equality in (6.1.3) is a special case of what is called *Cauchy's equation*, although, as we saw in Exercise 1.10, there is another equation, namely (1.10.2),that bears his name. However, the one here is much more significant.

Even more famous than the preceding is the following application of (6.1.4).

Theorem 6.1.2 *Continue in the setting of Theorem 6.1.1. Then for any continuous $f : \bar{G} \longrightarrow \mathbb{C}$ that is analytic on G and any $z \in G$*

$$f(z) = \frac{1}{i2\pi}\int_0^{2\pi} \frac{f\big(z_G(\varphi)\big)}{z_G(\varphi) - z} z'_G(\varphi)\,d\varphi. \tag{6.1.5}$$

Proof Apply (6.1.4) to the function $\zeta \rightsquigarrow \frac{f(\zeta)}{\zeta - z}$ to see that for small $r > 0$

$$\int_0^{2\pi} \frac{f\big(z_G(\varphi)\big)}{z_G(\varphi) - z} z'_G(\varphi)\,d\varphi = i\int_0^{2\pi} f\big(z + re^{i\varphi}\big)\,d\varphi,$$

and therefore, after letting $r \searrow 0$, that (6.1.5) holds. □

The equality in (6.1.5) is special case of the renowned *Cauchy integral formula*, a formula that is arguably the most profound expression of the remarkable properties of analytic functions.

Before moving to applications of these results, there is an important comment that should be made about the preceding. Namely, one should realize that (6.1.3) need not hold in general unless the function is analytic on the entire region. For example, if $z \in G$, then the function $\zeta \rightsquigarrow \frac{1}{\zeta - z}$ is analytic on $G\backslash\{z\}$ and yet (6.1.5) with $f = \mathbf{1}$ shows that

$$\int_0^{2\pi} \frac{1}{z_G(\varphi) - z} z_G'(\varphi)\, d\varphi = i2\pi. \tag{6.1.6}$$

One might think that the failure of (6.1.3) for this function is due to its having a singularity at z. However, the reason is more subtle. To wit,

$$\int_0^{2\pi} \frac{1}{(z_G(\varphi) - z)^n} z_G'(\varphi)\, d\varphi = 0 \quad \text{for } n \geq 2 \tag{6.1.7}$$

even though $\zeta \rightsquigarrow \frac{1}{(\zeta - z)^n}$ is more singular than $\zeta \rightsquigarrow \frac{1}{\zeta - z}$. To prove (6.1.7), note that $\partial_\zeta(\zeta - z)^{-n+1} = (1-n)(\zeta - z)^{-n}$, and therefore

$$(1-n)^{-1} \frac{d}{d\varphi} \big(z_G(\varphi) - z\big)^{-n+1} = \big(z_G(\varphi) - z\big)^{-n} z_G'(\varphi).$$

Hence, by the Fundamental Theorem of Calculus,

$$\int_0^{2\pi} \frac{1}{(z_G(\varphi) - z)^n} z_G'(\varphi)\, d\varphi = \frac{\Big(\big(z_G(b) - z\big)^{-n+1} - \big(z_G(a) - z\big)^{-n+1}\Big)}{1-n} = 0$$

since $z_G(a) = z_G(b)$. Of course, one might say that the same reasoning should work for $\zeta \rightsquigarrow (\zeta - z)^{-1}$ since $\partial_\zeta \log(\zeta - z) = (\zeta - z)^{-1}$. But unlike $\zeta \rightsquigarrow (\zeta - z)^{-n+1}$, log is not well defined on the whole of $\mathbb{C}\backslash\{z\}$ and must be restricted to a set like $\mathbb{C}\backslash(-\infty, 0]$ before it can be given an unambiguous definition. In particular, if one evaluates log along a path like z_G that has circled counterclockwise once around a point when it returns to its starting point, then the value of log when the path returns to its starting point will differ from its initial value by $i2\pi$, and this is the $i2\pi$ in (6.1.6). To see this, consider the path $\varphi \in [0, 2\pi] \longmapsto z(\varphi) \equiv z_G(\varphi) - c \in \mathbb{C}\backslash\{0\}$. This path is closed and circles the origin once counterclockwise. Using the definition of log in (2.2.2), one sees that

$$\log z(\varphi) = \log r(\varphi) + i \begin{cases} \varphi & \text{if } \varphi \in [0, \pi] \\ \varphi - 2\pi & \text{if } \varphi \in (\pi, 2\pi]. \end{cases}$$

Thus, since $r(2\pi) = r(0)$,

$$\int_0^{2\pi} \frac{1}{z_G(\varphi) - c} z_G'(\varphi)\, d\varphi = \int_0^{\pi} \frac{z'(\varphi)}{z(\varphi)}\, d\varphi + \lim_{\delta \searrow 0} \int_{\pi+\delta}^{2\pi-\delta} \frac{z'(\varphi)}{z(\varphi)}\, d\varphi$$
$$= \log \frac{r(\pi)}{r(0)} + i\pi + \lim_{\delta \searrow 0} \Big(\log \frac{r(2\pi - \delta)}{r(\pi + \delta)} + i\pi \Big) = i2\pi.$$

Another way to prove (6.1.6) is to use (6.1.7) to see that

$$\partial_z \int_0^{2\pi} \frac{1}{z_G(\varphi) - z} z_G'(\varphi)\, d\varphi = \int_0^{2\pi} \frac{1}{(z_G(\varphi) - z)^2} z'(\varphi)\, d\varphi = 0.$$

Thus if (6.1.6) holds for one $z \in G$, then it holds for all $z \in G$, and, because it is obvious when $z = c$, this completes the proof.

6.2 A Few Theoretical Applications and an Extension

There are so many applications of (6.1.3) and (6.1.5) that one is at a loss when trying to choose among them. One of the most famous is the *fundamental theorem of algebra* which says that if p is a polynomial of degree $n \geq 1$ then it has a root (i.e., a $z \in \mathbb{C}$ such that $p(z) = 0$). To prove this, first observe that, by applying (6.1.5) with $G = D(z, R)$ and $z_G(\varphi) = z + Re^{i\varphi}$, one has the *mean value theorem*

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} f\big(z + Re^{i\varphi}\big)\, d\varphi \tag{6.2.1}$$

for any f that is continuous on $\overline{D(z, R)}$ and analytic on $D(z, R)$. Now suppose that p is an nth order polynomial for some $n \geq 1$. Then $|p(z)| \longrightarrow \infty$ as $|z| \to \infty$. Thus, if p never vanished, $f = \frac{1}{p}$ would be an analytic function on $\mathbb{C}$ that tends to 0 as $|z| \to \infty$. But then, because (6.2.1) holds for all $R > 0$, after letting $R \to \infty$ one would get the contradiction that $f(z) = 0$ for all $z \in \mathbb{C}$, and so there must be a $z \in \mathbb{C}$ for which $p(z) = 0$. In conjunction with Lemma 2.2.3, this means that every nth order polynomial p is equal to $b \prod_{j=1}^{\ell} (z - a_j)^{d_j}$, where b is the coefficient of z^n, $a_1, \dots, a_\ell$ are the distinct roots of p, and d_j is the *multiplicity* of a_j. That is, d_j is the order to which p vanishes at a_j in the sense that $\lim_{z \to a_j} \frac{p(z)}{(z - a_j)^{d_j}} \in \mathbb{C} \backslash \{0\}$.

As a by-product of the preceding argument, we know that if f is analytic on $\mathbb{C}$ and tends to 0 at infinity, then f is identically 0. A more refined version of this property is the following.

Theorem 6.2.1 *If f is a bounded, analytic function on $\mathbb{C}$, then f is constant.*

Proof Starting from (6.2.1), one sees that

$$\pi R^2 f(z) = 2\pi \int_0^R r f(z)\,dr = \int_0^R r\left(\int_0^{2\pi} f\big(z+re^{i\varphi}\big)\,d\varphi\right)dr,$$

which, by Theorem 5.6.2, leads to

$$f(z) = \frac{1}{\pi R^2}\int_{\overline{D(z,R)}} f(\zeta)\,d\xi d\eta \tag{6.2.2}$$

for any f that is continuous on $\overline{D(0,R)}$ and analytic on $D(0,R)$. Now assume that f is a function that is analytic on the whole of $\mathbb{C}$. Then for any z_1 and z_2 and any $R > |z_2 - z_1|$

$$f(z_2) - f(z_1) = \frac{1}{\pi R^2}\left(\int_{\overline{D(z_2,R)}\setminus D(z_1,R)} f(\zeta)\,d\xi d\eta - \int_{\overline{D(z_1,R)}\setminus D(z_2,R)} f(\zeta)\,d\xi d\eta\right).$$

Because

$$\begin{aligned}\operatorname{vol}\big(\overline{D(z_1,R)}\setminus D(z_2,R)\big) &= \operatorname{vol}\big(\overline{D(z_2,R)}\setminus D(z_1,R)\big)\\ &\le \operatorname{vol}\big(D(z_2,R+|z_2-z_1|)\setminus D(z_1,R)\big) = \pi\big(2R|z_1-z_2| + |z_1-z_2|^2\big),\end{aligned}$$

$|f(z_2) - f(z_1)| \le \|f\|_{\mathbb{C}}\frac{6|z_2-z_1|}{R}$. Now let $R \nearrow \infty$, and conclude that $f(z_2) = f(z_1)$ if $\|f\|_{\mathbb{C}} < \infty$. □

The fact that the only bounded, analytic functions on $\mathbb{C}$ are constant is known as *Liouville's theorem*.

Recall that our original examples in Sect. 2.3 of analytic functions were power series. The following remarkable consequence of (6.1.5) shows that was inevitable.

Theorem 6.2.2 *Suppose that f is analytic on a non-empty open set $G \subseteq \mathbb{C}$. If $\overline{D(\zeta,R)} \subseteq G$, then f is given on $D(\zeta,R)$ by the power series*

$$\begin{gathered} f(z) = \sum_{m=0}^{\infty} c_m (z-\zeta)^m \quad \textit{for } z \in D(\zeta,R)\\ \textit{where } c_m = \frac{1}{2\pi R^m}\int_0^{2\pi} f(\zeta + Re^{i\varphi})e^{-im\varphi}\,d\varphi. \end{gathered} \tag{6.2.3}$$

In particular, if G is connected and f as well as all its derivatives vanish at a point $\zeta \in G$, then f is identically 0 on G.

Proof By applying (6.1.5), we know that

$$f(z) = \frac{R}{2\pi}\int_0^{2\pi} \frac{f(\zeta + Re^{i\varphi})}{Re^{i\varphi} - (z-\zeta)} e^{i\varphi}\, d\varphi \quad \text{for } z \in D(\zeta, R),$$

and, because $|z-\zeta| < R$ and therefore

$$\frac{Re^{i\varphi}}{Re^{i\varphi} - (z-\zeta)} = \left(1 - \frac{(z-\zeta)e^{-i\varphi}}{R}\right)^{-1} = \sum_{m=0}^{\infty} \frac{(z-\zeta)^m e^{-im\varphi}}{R^m},$$

where the convergence is absolute and uniform for $\varphi \in [0, 2\pi]$, (6.2.3) follows.

To prove the final assertion, let $H = \{z \in G : f^{(m)}(z) = 0 \text{ for all } m \geq 0\}$. Clearly, $G\backslash H$ is open. On the other hand, if $\zeta \in H$ and $R > 0$ is chosen so that $\overline{D(\zeta, R)} \subseteq G$, then (6.2.3) says that f vanishes on $D(\zeta, R)$ and therefore that $D(\zeta, R) \subseteq H$. Hence, both H and $G\backslash H$ are open. Since G is connected, it follows that either $H = \emptyset$ or $H = G$. □

What is startling about this result is how much it reveals about the structure of an analytic function. Recall that f is analytic as soon as it is *once* continuously differentiable as a function of a complex variable. As a consequence of Theorem 6.2.2, we now know that if a function is once continuously differentiable as a function of a complex variable, then it is not only an infinitely differentiable function of that variable but is also given by a power series on any disk whose closure is contained in the set where it is once continuously differentiable. In addition, (6.2.3) gives quantitative information about the derivatives of f. Namely,

$$f^{(m)}(\zeta) = m!c_m = \frac{m!}{2\pi R^m}\int_0^{2\pi} f(\zeta + Re^{i\varphi})e^{-im\varphi}\, d\varphi. \tag{6.2.4}$$

Notice that this gives a second proof Liouville's theorem. Indeed, if f is bounded and analytic on $\mathbb{C}$, then (6.2.4) shows that f' vanishes everywhere and therefore, by Lemma 2.3.1, f is constant.

To demonstrate one way in which these results get applied, recall the sequence of numbers $\{b_\ell : \ell \geq 0\}$ introduced in (3.3.4). As we saw there, $|b_\ell| \leq \alpha^\ell$ where α is the unique element of $(0, 1)$ for which $e^{\frac{1}{\alpha}} = 1 + \frac{2}{\alpha}$, and in (3.4.8) we gave an expression for b_ℓ from which it is easy to show that $\overline{\lim}_{\ell\to\infty} |b_\ell|^{\frac{1}{\ell}} = \frac{1}{2\pi}$. We will now show how to arrive at the second of these as a consequence of the first combined with Theorem 6.2.2. For this purpose, consider the function $f(z) = \sum_{\ell=0}^{\infty} b_{\ell+1} z^\ell$, which we know is an analytic function on $D(0, \alpha^{-1})$. Next, using the recursion relation in (3.3.4) for the b_ℓ's, observe that for $z \in D(0, \alpha^{-1})\backslash\{0\}$,

$$f(z) \equiv \sum_{k=0}^{\infty} z^k \left(\sum_{\ell=k}^{\infty} \frac{(-1)^k b_{\ell-k}}{(k+2)!} z^{\ell-k}\right) = \big(1 + zf(z)\big)\frac{e^{-z} - 1 + z}{z^2},$$

and therefore that $f(z) = \frac{1-e^z+ze^z}{z(e^z-1)}$. Thus, if $\varphi(z) = \sum_{m=0}^{\infty} \frac{m+1}{(m+2)!} z^m$ and $\psi(z) = \sum_{m=0}^{\infty} \frac{1}{(m+1)!} z^m$, then $f(z) = \frac{\varphi(z)}{\psi(z)}$ for $z \in D(0, \alpha^{-1})$. Because φ and ψ are both analytic on the whole of $\mathbb{C}$, their ratio is analytic on the set where $\psi \neq 0$. Furthermore, since $\psi(0) \neq 0$ and $\psi(z) = \frac{e^z-1}{z}$ for $z \neq 0$, $\psi(z) = 0$ if and only if $z = i2\pi n$ for some $n \in \mathbb{Z}\setminus\{0\}$, which, because $\varphi(i2\pi) \neq 0$, implies that $D(0, 2\pi)$ is the largest disk centered at 0 on which their ratio is analytic. But this means that the power series defining f is absolutely convergent when $|z| < 2\pi$ and divergent when $|z| > 2\pi$. In other words, its radius of convergence must be 2π, and therefore, by Lemma 2.2.1, $\overline{\lim}_{\ell\to\infty} |b_\ell|^{\frac{1}{\ell}} = (2\pi)^{-1}$.

Here are a few more results that show just how special analytic functions are.

Lemma 6.2.3 *Suppose that $f : \overline{D(a, R)}\setminus\{a\} \longrightarrow \mathbb{C}$ is a continuous function which is analytic on $D(a, R)\setminus\{a\}$. Then, for each $0 < r < R$ and $z \in D(a, R)\setminus\overline{D(a, r)}$*

$$f(z) = \frac{R}{2\pi} \int_0^{2\pi} \frac{f\big(a + Re^{i\varphi}\big)e^{i\varphi}}{a + Re^{i\varphi} - z}\, d\varphi - \frac{r}{2\pi} \int_0^{2\pi} \frac{f\big(a + re^{i\varphi}\big)e^{i\varphi}}{a + re^{i\varphi} - z}\, d\varphi.$$

Proof Without loss in generality, we will assume that $a = 0$.

Let $0 < r < R$ and $z \in D(0, R)\setminus\overline{D(0, r)}$ be given, and choose $\rho > 0$ so that $\overline{D(z, \rho)} \subseteq D(0, R)\setminus\overline{D(0, r)}$. By (6.1.4) applied to $\zeta \rightsquigarrow \frac{f(\zeta)}{\zeta - z}$, we see that

$$R \int_0^{2\pi} \frac{f(Re^{i\varphi})e^{i\varphi}}{Re^{i\varphi} - z}\, d\varphi = r \int_0^{2\pi} \frac{f(re^{i\varphi})e^{i\varphi}}{re^{i\varphi} - z}\, d\varphi + \int_0^{2\pi} f(z + \rho e^{i\varphi})\, d\varphi,$$

and therefore, after letting $\rho \searrow 0$, we have that

$$f(z) = \frac{R}{2\pi} \int_0^{2\pi} \frac{f\big(Re^{i\varphi}\big)e^{i\varphi}}{Re^{i\varphi} - z}\, d\varphi - \frac{r}{2\pi} \int_0^{2\pi} \frac{f\big(re^{i\varphi}\big)e^{i\varphi}}{re^{i\varphi} - z}\, d\varphi. \qquad \square$$

There are two interesting corollaries of Lemma 6.2.3.

Theorem 6.2.4 *Referring to Lemma 6.2.3, define*

$$c_m = \frac{1}{2\pi R^m} \int_0^{2\pi} f(a + Re^{i\varphi})e^{-im\varphi}\, d\varphi \quad \text{for } m \in \mathbb{Z}.$$

Then

$$f(z) = \sum_{m=-\infty}^{\infty} c_m (z - a)^m \quad \text{for } z \in D(a, R)\setminus\{a\},$$

where the series converges absolutely and uniformly for z in compact subsets of $D(a, R)\setminus\{a\}$.

Proof We can and will assume that $a = 0$.

Let $z \in D(0, R)\backslash\{0\}$ and $r > 0$ satisfying $\overline{D(z, r)} \subseteq D(0, R)\backslash\{0\}$ be given. Look at the first term on the right hand side of the equation in Lemma 6.2.3, and observe that it can be re-written as

$$\frac{1}{2\pi}\int_0^{2\pi} \frac{f(Re^{i\varphi})}{1 - (Re^{i\varphi})^{-1}z}\, d\varphi = \frac{1}{2\pi}\sum_{m=0}^{\infty} c_m z^m,$$

where the convergence is absolute and uniform for z in compact subsets of $D(0, R)$. The second term can be written as

$$\frac{r}{2\pi z}\int_0^{2\pi} \frac{f(re^{i\varphi})e^{i\varphi}}{1 - re^{i\varphi}z^{-1}}\, d\varphi = \frac{1}{2\pi}\sum_{m=1}^{\infty}\left(\frac{r}{z}\right)^m \int_0^{2\pi} f(re^{i\varphi})e^{im\varphi}\, d\varphi,$$

where the series converges absolutely and uniformly on compact subsets of $D(0, R)\backslash \overline{D(0, r)}$. Finally, apply (6.1.4) to the analytic function $\zeta \rightsquigarrow f(\zeta)\zeta^{m-1}$ on $D(0, R)\backslash \overline{D(0, r)}$ to see that

$$r^m \int_0^{2\pi} f(re^{i\varphi})e^{im\varphi}\, d\varphi = R^m \int_0^{2\pi} f(Re^{i\varphi})e^{im\varphi}\, d\varphi$$

and therefore that the asserted expression for f holds on $D(0, R)\backslash\overline{D(0, r)}$. In addition, since the series converges absolutely and uniformly for z in compact subsets of $D(0, R)\backslash\overline{D(0, r)}$ for each $0 < r < R$, this completes the proof. □

Theorem 6.2.5 *Let $G \subseteq \mathbb{C}$ be open and $a \in G$. If f is analytic on $G\backslash\{a\}$ and* $\lim_{r\searrow 0} r\|f\|_{\mathbb{S}^1(a,r)} = 0$, *then there is a unique extension of f to G as an analytic function.*

Proof The uniqueness poses no problem. In addition, it suffices for us to show that f admits an extension as an analytic function on some disk centered at a.

Choose $R > 0$ so that $\overline{D(a, R)} \subseteq G$, and let $z \in D(a, R)\backslash\{a\}$ and $r > 0$ satisfying $\overline{D(z, r)} \subseteq D(a, R)\backslash\{a\}$ be given. Applying Lemma 6.2.3 and then letting $r \searrow 0$, we arrive at

$$f(z) = \frac{R}{2\pi}\int_0^{2\pi} \frac{f\left(a + Re^{i\varphi}\right)}{a + Re^{i\varphi} - z} e^{i\varphi}\, d\varphi.$$

Hence, the right hand side gives the analytic extension of f to $D(a, R)$. □

Theorem 6.2.6 *Suppose $\{f_n : n \geq 1\}$ is a sequence of analytic functions on an open set G and that f is a function on G to which $\{f_n : n \geq 1\}$ converges uniformly on compact subsets of G. Then f is analytic on G.*

Proof Given $c \in G$, choose $r > 0$ so that $\overline{D(c, r)} \subseteq G$. All that we have to show is that f is analytic on $D(c, r)$. But $f(z)$ equals

$$\lim_{n\to\infty} f_n(z) = \lim_{n\to\infty} \frac{1}{2\pi}\int_0^{2\pi} \frac{f_n(c+re^{i\varphi})e^{i\varphi}}{c+r^{i\varphi}-z}\,d\varphi = \frac{1}{2\pi}\int_0^{2\pi} \frac{f(c+re^{i\varphi})e^{i\varphi}}{c+r^{i\varphi}-z}\,d\varphi$$

for all $z \in D(c,r)$, and, since the final expression is an analytic function of $z \in D(c,r)$, this proves that f is analytic on $D(c,r)$. □

Related to the discussion leading to (6.1.7) and (6.1.6) is the following. As we saw there, (6.1.3) is obvious in the case when f is the derivative of a function. Thus the following important fact provides a way to prove more general versions of (6.1.3) and (6.1.5).

Theorem 6.2.7 *If f is analytic on the star shaped region G, then there exists an analytic function F on G such that $f = F'$. Furthermore, F is uniquely determined up to an additive constant.*

Proof To prove the uniqueness assertion, suppose that F and $\tilde{F}$ are analytic functions on G such that $F' = f = \tilde{F}'$ there, and set $g = F - \tilde{F}$. Then g is analytic and $g' = 0$ on G. Since (cf. Exercise 4.5) G is connected, Lemma 2.3.1 says that g is constant.

Not surprisingly, the existence assertion is closely related to Corollary 5.7.3. Indeed, write $f = u + iv$, and take $\mathbf{F} = (u, -v)$ and $\tilde{\mathbf{F}} = (v, u)$ as we did in the derivation of (6.1.1). Observe that, by (6.0.1), both $\mathbf{F}$ and $\tilde{\mathbf{F}}$ are exact. Thus, by Corollary 5.7.3, there exist continuously differentiable functions U and V on G such that $u = \partial_x U$, $v = -\partial_y U$, $v = \partial_x V$, and $u = \partial_y V$. In particular, $\partial_x U = u = \partial_y V$ and $\partial_y U = -v = -\partial_x V$, and so U and V satisfy the Cauchy–Riemann equations in (6.0.1), which means that $F \equiv U + iV$ is analytic. Furthermore, $\partial_x F = u + iv = f$, and therefore $F' = f$. □

We will say that a path $z : [a, b] \longrightarrow \mathbb{C}$ is *closed* if $z(b) = z(a)$ and that it is *piecewise smooth* if it is continuous and there exists a bounded function $z' : [a, b] \longrightarrow \mathbb{C}$ that has at most a finite number of discontinuities and for which $z'(t)$ is the derivative of $t \rightsquigarrow z(t)$ at all its points of continuity.

Corollary 6.2.8 *Suppose that f is an analytic function on a star shaped region G, and let $z : [a, b] \longrightarrow G$ be a closed, piecewise smooth path. Then (6.1.3) holds with $t \rightsquigarrow z_G(t)$ replaced by $t \rightsquigarrow z(t)$. Furthermore, if $z \in G\setminus\{z(t) : t \in [a, b]\}$, then*

$$\int_a^b \frac{f\big(z(t)\big)}{z(t)-z} z'(t)\,dt = \left(\int_a^b \frac{z'(t)}{z(t)-z}\,dt\right) f(z). \tag{6.2.5}$$

Proof By Theorem 6.2.7, there exists an analytic function F on G for which $f = F'$. Thus $\frac{d}{dt}F\big(z(t)\big) = f\big(z(t)\big)z'(t)$ for all but a finite number of $t \in (a, b)$, and therefore, by the Fundamental Theorem of Calculus,

$$\int_a^b f\big(z(t)\big)z'(t)\,dt = F\big(z(b)\big) - F\big(z(a)\big) = 0.$$

Next, given $z \in G\backslash\{z(t) : t \in [a,b]\}$, define

$$\zeta \in G\backslash\{z\} \longmapsto g(\zeta) = \frac{f(\zeta) - f(z)}{\zeta - z} \in \mathbb{C}.$$

Then g is analytic and $\lim_{\zeta \to z} g(\zeta) = f'(z)$. Hence, by Theorem 6.2.5, g can be extended to G as an analytic function, and therefore, by the preceding,

$$\int_a^b \frac{f\big(z(t)\big)}{z(t) - z} z'(t)\,dt - f(z) \int_a^b \frac{z'(t)}{z(t) - z}\,dt = \int_a^b \frac{f\big(z(t)\big) - f(z)}{z(t) - z} z'(t)\,dt = 0.$$

□

The equation in (6.2.5) represents a significant extension of *Cauchy's formula*, although, as stated, it does not quite cover (6.1.5). However, it is easy to recover (6.1.5) from (6.2.5) by applying the same approximation procedure as we used to complete the proof of Theorem 6.1.1.

Referring to Corollary 6.2.8, the quantity $\int_a^b \frac{z'(t)}{z(t)-z}\,dt$ has an important geometric interpretation. When $t \rightsquigarrow z(t)$ parameterizes the boundary of a piecewise smooth star shaped region, (6.1.6) says that it is equal to $i2\pi$, and, as the discussion preceding (6.1.6) indicates, that is because $t \rightsquigarrow z(t)$ circles z exactly one time. In order to deal with more general paths, we will need the following lemma.

Lemma 6.2.9 *If $z \in \mathbb{C}$ and $t \in [a,b] \longmapsto z(t) \in \mathbb{C}\backslash\{z\}$ is continuous, then there exists a unique continuous map $\theta : [a,b] \longrightarrow \mathbb{R}$ such that $\theta(a) = 0$ and*

$$\frac{z(t) - z}{z(a) - z} = \frac{|z(t) - z|}{|z(a) - z|} e^{i\theta(t)} \quad \text{for } t \in [a,b].$$

Furthermore, if $t \rightsquigarrow z(t)$ is piecewise smooth, then

$$\theta(t) = -i \int_a^t \frac{z'(s)}{z(s) - z}\,ds \quad \text{for } t \in [a,b].$$

In particular, if, in addition, $z(b) = z(a)$, then

$$\frac{1}{i2\pi} \int_a^b \frac{z'(s)}{z(s) - z}\,ds \in \mathbb{Z}.$$

(See Exercise 6.9 for another derivation of this last fact.)

Proof Without loss in generality, we will assume that $z = 0$ and $z(a) = 1$.

Because $t \rightsquigarrow z(t)$ is continuous and never vanishes, there exist $n \geq 1$ and $a = t_0 < \cdots < t_n = b$ such that $\left|\frac{z(t)}{z(t_{m-1})} - 1\right| < 1$ for $1 \leq m \leq n$ and $t \in [t_{m-1}, t_m]$. Now set $\ell(a) = 0$ and $\ell(t) - \ell(t_{m-1}) = \log \frac{z(t)}{z(t_{m-1})}$ for $1 \leq m \leq n$ and $t \in [t_{m-1}, t_m]$. Then ℓ is continuous and, for each $1 \leq m \leq n$, $z(t) = z(t_{m-1}) e^{\ell(t) - \ell(t_{m-1})}$ for $t \in [t_{m-1}, t_m]$.

Hence, by induction on m, $z(t) = e^{\ell(t)}$ for all $t \in [a,b]$. Furthermore, if $t \rightsquigarrow z(t)$ is piecewise smooth, then $\ell'(t) = \frac{z'(t)}{z(t)}$ at all but at most a finite number of $t \in (a,b)$, and therefore $\ell(t) = \int_a^t \frac{z'(s)}{z(s)}\,ds$. Thus, if we take $\theta(t) = -i\ell(t)$, then $t \rightsquigarrow \theta(t)$ will have all the required properties.

Finally, to prove the uniqueness assertion, suppose that $t \rightsquigarrow \theta_1(t)$ and $t \rightsquigarrow \theta_2(t)$ were two such functions. Then $e^{i(\theta_2(t)-\theta_1(t))} = 1$ for all $t \in [a,b]$, and so $t \rightsquigarrow \theta_2(t)-\theta_1(t)$ would be a continuous function with values in $\{2\pi n : n \in \mathbb{Z}^+\}$, which is possible (cf. the proof of (2.3.2)) only if it is constant. Thus, since $\theta_1(a) = 0 = \theta_2(a)$, $\theta_1(t) = \theta_2(t)$ for all $t \in [a,b]$. □

The function $t \in [a,b] \longmapsto \theta(t) \in \mathbb{R}$ gives the one and only continuous choice of the angle in the polar representation of $\frac{z(t)-z}{z(a)-z}$ under the condition that $\theta(a) = 0$. Given $a \le s_1 < s_2 \le b$, we say that the path circles z once during the time interval $[s_1, s_2]$ if $|\theta(s_2) - \theta(s_1)| = 2\pi$ and $|\theta(t) - \theta(s_1)| < 2\pi$ for $t \in [s_1, s_2)$. Further, we say it circled counterclockwise or clockwise depending on whether $\theta(s_2) - \theta(s_1)$ is 2π or -2π. Thus, if $t \rightsquigarrow z(t)$ is closed and one were to put a rod perpendicular to the plane at z and run a string along the path, then $\frac{\theta(b)-\theta(a)}{2\pi}$ would the number of times that the string winds around the rod after one pulls out all the slack, and for this reason it is called the *winding number* of the path $t \rightsquigarrow z(t)$ around z. Hence, when $t \rightsquigarrow z(t)$ is piecewise smooth, (6.2.5) is saying that

$$\int_a^b \frac{f\big(z(t)\big)}{z(t)-z} z'(t)\,dt = i2\pi f(z)(\text{the winding number of } t \rightsquigarrow z(t) \text{ around } z).$$

In the case when $t \rightsquigarrow z(t)$ is the parameterization of the boundary of a piecewise smooth star shaped region, (6.1.6) says that its winding number around every point in the region is 1.

6.3 Computational Applications of Cauchy's Formula

The proof of Corollary 6.2.8 can be viewed as an example of a ubiquitous and powerful technique, known as the *residue calculus*, for computing integrals.

Suppose that $f : \overline{D(a,R)}\backslash\{a\} \longrightarrow \mathbb{C}$ is a continuous function that is analytic on $D(a,R)\backslash\{a\}$. Then

$$\mathrm{Res}_a(f) \equiv \frac{R}{2\pi}\int_0^{2\pi} f(a + Re^{i\varphi})e^{i\varphi}\,d\varphi$$

is called the *residue* of f at a. Notice that, because, by (6.1.4),

$$R_1\int_0^{2\pi} f(a + R_1e^{i\varphi})e^{i\varphi}\,d\varphi = R_2\int_0^{2\pi} f(a + R_2e^{i\varphi})e^{i\varphi}\,d\varphi$$

if $0 < R_1 < R_2$ and $f : \overline{D(a, R_2)}\setminus\{a\} \longrightarrow \mathbb{C}$ is a continuous function that is analytic on $D(a, R_2)\setminus\{a\}$, $\mathrm{Res}_a(f)$ really depends only on f and a, and not on R.

The computation of residues is difficult in general. However there is a situation in which it is relatively easy. Namely, if $n \geq 1$ and $(z-a)^n f(z)$ is bounded on $\overline{D(a, R)}\setminus\{a\}$, then, by Theorem 6.2.5, there is an analytic function g on $D(a, R)$ such that $g(z) = (z-a)^n f(z)$ for $z \neq a$. Hence

$$f(z) = \sum_{m=0}^{\infty} \frac{g^{(m+n)}(a)}{(m+n)!}(z-a)^m + \sum_{m=1}^{n} \frac{g^{(n-m)}(a)}{(n-m)!}(z-a)^{-m},$$

and so, since

$$\int_0^{2\pi} e^{i(m+1)\varphi}\,d\varphi = 0 \quad \text{unless } m = -1,$$

we have that

$$R\int_0^{2\pi} f(a+Re^{i\varphi})e^{i\varphi}\,d\varphi = \frac{2\pi g^{(n-1)}(a)}{(n-1)!}.$$

Equivalently,

$$\begin{aligned}&\mathrm{Res}_a(f) = \tfrac{g^{(n-1)}(a)}{(n-1)!} \quad \text{if } g(z) = (z-a)^n f(z)\\ &\text{is bounded on } D(a, R)\setminus\{a\} \text{ for some } R > 0.\end{aligned} \tag{6.3.1}$$

The following theorem gives the essential facts on which the residue calculus is based.

Theorem 6.3.1 *Let G be a star shaped region and $z_1, \dots, z_\ell$ distinct points in G. If $f : G\setminus\{z_1, \dots, z_\ell\} \longrightarrow \mathbb{C}$ is an analytic function and $z : [a, b] \longrightarrow G\setminus\{z_1, \dots, z_\ell\}$ is a closed piecewise smooth path, then*

$$\int_a^b f\big(z(t)\big)z'(t)\,dt = i2\pi\sum_{k=1}^{\ell} N(z_k)\mathrm{Res}_{z_k}(f),$$

where

$$N(z_k) = \frac{1}{i2\pi}\int_a^b \frac{z'(t)}{z(t)-z_k}\,dt$$

is the winding number of $t \rightsquigarrow z(t)$ around z_k.

Proof Choose $R_1, \dots, R_\ell > 0$ so that $\overline{D(z_k, R_k)} \subseteq G\setminus\{z_j : j \neq k\}$ for $1 \leq k \leq \ell$, and set

$$c_{k,m} = \frac{1}{2\pi R_k^m}\int_0^{2\pi} f(z_k + R_k e^{i\varphi})e^{-im\varphi}\,d\varphi.$$

By Theorem 6.2.4, we know that the series $\sum_{m=-\infty}^{\infty} c_{k,m}(z-z_k)^m$ converges absolutely and uniformly for z in compact subsets of $D(z_k, R_k)\setminus\{z_k\}$, and therefore that $H_k(z) \equiv \sum_{m=1}^{\infty} c_{k,-m}(z-z_k)^{-m}$ converges absolutely and uniformly for z in compact subsets of $\mathbb{C}\setminus\{z_k\}$. Furthermore, by that same theorem, $g(z) \equiv f(z) - \sum_{k=1}^{\ell} H_k(z)$ on $G\setminus\{z_1, \dots, z_\ell\}$ admits an extension as an analytic function on G. Hence, by the first part of Corollary 6.2.8,

$$\int_a^b f\big(z(t)\big)z'(t)\,dt = \sum_{k=1}^{\ell} \int_a^b H_k\big(z(t)\big)z'(t)\,dt.$$

Finally, for each $1 \le k \le \ell$,

$$\begin{aligned}\int_a^b H_k\big(z(t)\big)z'(t)\,dt &= \sum_{m=1}^{\infty} c_{k,-m} \int_a^b \frac{z'(t)}{\big(z(t)-z_k\big)^m}\,dt \\ &= c_{k,-1}\int_a^b \frac{z'(t)}{z(t)-z_k}\,dt = i2\pi N(z_k)\mathrm{Res}_{z_k}(f),\end{aligned}$$

since $c_{k,-1} = \mathrm{Res}_{z_k}(f)$ and

$$\int_a^b \frac{z'(t)}{\big(z(t)-z_k\big)^m}\,dt = \frac{1}{(1-m)}\big(z(t)-z_k\big)^{1-m}\Big|_{t=a}^{b} = 0$$

if $m \ge 2$. □

In order to present our initial application of Theorem 6.3.1, we need to make some preparations. When confronted by an integral like $\int_0^1 \frac{1}{x^2+5x+6}\,dx$, one factors x^2+5x+6 into the product $(x+2)(x+3)$ and uses this to write

$$\int_0^1 \frac{1}{x^2+5x+6}\,dx = \int_0^1 \frac{1}{x+2}\,dx - \int_0^1 \frac{1}{x+3}\,dx = \log\tfrac{3}{2} - \log\tfrac{4}{3} = \log\tfrac{9}{8}.$$

The crucial step here, which is known as the method of *partial fractions*, is the decomposition of $\frac{1}{x^2+5x+6}$ into the difference between $\frac{1}{x+2}$ and $\frac{1}{x+3}$. The following theorem shows that such a decomposition exists in great generality.

Theorem 6.3.2 *Suppose that q is a polynomial of order $n \ge 1$ and that $b_1, \dots, b_\ell$ are its distinct roots, and let p be a polynomial of order $m \ge 0$ satisfying $p(b_k) \ne 0$ for $1 \le k \le \ell$. Then there exist unique polynomials $P_0, \dots, P_\ell$ such that P_0 is of order $(n-m)^+$, P_k vanishes at 0 and has order equal to the multiplicity of b_k for each $1 \le k \le \ell$, and*

$$\frac{p(z)}{q(z)} = P_0(z) + \sum_{k=1}^{\ell} P_k\left(\frac{1}{z-b_k}\right) \quad \text{for } z \in \mathbb{C}\setminus\{b_1, \dots, b_\ell\}.$$

Proof By long division for polynomials, $\frac{p(z)}{q(z)} = P_0(z) + R_0(z)$, where P_0 is a polynomial of the specified sort and R_0 is the ratio of two polynomials for which the polynomial in the denominator has order larger than that in the numerator. Next, given $1 \le k \le \ell$, consider the function

$$z \rightsquigarrow \frac{p\big(b_k + \frac{1}{z}\big)}{q\big(b_k + \frac{1}{z}\big)}.$$

This function is again the ratio of two polynomials, and as such can be written as $P_k(z) + R_k(z)$, where P_k is a polynomial and R_k is the ratio of two polynomials for which the order of the one in the denominator is greater than that of the order of the one in the numerator. Moreover, if d_k is the multiplicity of b_k and c is the coefficient of z^n in q, then

$$\frac{P_k(z)}{z^{d_k}} = \frac{p\big(b_k + \frac{1}{z}\big)}{z^{d_k} q\big(b_k + \frac{1}{z}\big)} - \frac{R_k(z)}{z^{d_k}} \longrightarrow \frac{p(b_k)}{c\prod_{j \ne k}(b_k - b_j)^{d_j}} \in \mathbb{C}\backslash\{0\}$$

as $|z| \to \infty$, and therefore P_k has order d_k. Now consider the function

$$f(z) = R_0(z) - \sum_{j=1}^{\ell} P_j\left(\frac{1}{z - b_j}\right) \quad \text{for } z \in \mathbb{C}\backslash\{b_1, \dots, b_\ell\}.$$

Obviously f is analytic. In addition, as $|z| \to \infty$, $f(z)$ tends to

$$-\sum_{j=1}^{\ell} P_j(0) \in \mathbb{C}.$$

For $1 \le k \le \ell$,

$$\begin{aligned} f\big(b_k + \tfrac{1}{z}\big) &= R_k(z) - P_0\big(b_k + \tfrac{1}{z}\big) - \sum_{j \ne k} P_j\left(\frac{1}{\frac{1}{z} + b_k - b_j}\right) \\ &\longrightarrow -P_0(b_k) - \sum_{j \ne k} P_j\left(\frac{1}{b_k - b_j}\right) \in \mathbb{C} \end{aligned}$$

as $|z| \to \infty$. Hence f is bounded, and therefore, by Theorems 6.2.5 and 6.2.1, f must be identically equal to some constant b_0. Finally, for each $1 \le k \le \ell$, replace P_k by $P_k - P_k(0)$, and then replace P_0 by $P_0 + b_0 + \sum_{k=1}^{\ell} P_k(0)$.

Turning to the question of uniqueness, write $P_k(z) = \sum_{j=1}^{d_k} a_{k,j} z^j$ for $1 \le k \le \ell$. Then

$$a_{k,d_k} = \lim_{z \to b_k} (z - b_k)^{d_k} \frac{p(z)}{q(z)}$$

and, for $1 \le j < d_k$,

$$a_{k,j} = \lim_{z \to b_k} (z - b_k)^j \left(\frac{p(z)}{q(z)} - \sum_{j < j' \le d_k} \frac{a_{k,j'}}{(z - b_k)^{j'}} \right).$$

Hence $P_1, \dots, P_k$ are uniquely determined, and therefore so is P_0. □

In general the computation of the P_k's in Theorem 6.3.2 is quite onerous. However, there is an important case in which it is easy. Namely, suppose that q is a polynomial of order $n \ge 1$ all of whose roots are *simple* in the sense that they are of multiplicity 1. Then (cf. Exercise 6.4 below for an elementary algebraic proof)

$$\frac{1}{q(z)} = \sum_{k=1}^{n} \frac{1}{q'(b_k)(z - b_k)}, \tag{6.3.2}$$

where $b_1, \dots, b_n$ are the roots of q. Indeed, here $P_0 = 0$, $q'(b_k)$ equals the coefficient of z^n times $\prod_{j \ne k}(b_k - b_j)$ and is therefore not 0, and $P_k(z) = a_k z$ where $a_k = \lim_{z \to b_k} \frac{z - b_k}{q(z)} = \frac{1}{q'(b_k)}$ for $1 \le k \le \ell$.

Corollary 6.3.3 *Let q be a polynomial of order $n \ge 2$ with simple roots $b_1, \dots, b_n$. Then for any polynomial p of order less than or equal to $n - 2$,*

$$\sum_{k=1}^{n} \frac{p(b_k)}{q'(b_k)} = 0.$$

Moreover, if none of the b_k's is real and $K_\pm$ is the set of $1 \le k \le n$ for which the imaginary part of $\pm b_k$ is positive, then

$$\int_{\mathbb{R}} \frac{p(x)}{q(x)}\, dx = i2\pi \sum_{k \in K_+} \frac{p(b_k)}{q'(b_k)} = -i2\pi \sum_{k \in K_-} \frac{p(b_k)}{q'(b_k)}.$$

In particular, if either K_+ or K_- is empty, then $\int_{\mathbb{R}} \frac{p(x)}{q(x)}\, dx = 0$.

Proof Let $R > \max\{|b_j| : 1 \le j \le n\}$. Then, by (6.3.2), (6.3.1), and Theorem 6.3.1,

$$R \int_0^{2\pi} \frac{p\big(Re^{i\varphi}\big)}{q\big(Re^{i\varphi}\big)} e^{i\varphi}\, d\varphi = 2\pi \sum_{j=1}^{n} \frac{p(b_j)}{q'(b_j)}.$$

Since

$$\lim_{R \to \infty} R \sup_{\varphi \in [0, 2\pi]} \left| \frac{p\big(Re^{i\varphi}\big)}{q\big(Re^{i\varphi}\big)} \right| = 0,$$

the first assertion follows.

Next assume that none of the b_k's is real. For $R > \max\{|b_k| : 1 \le k \le n\}$, set $G_R = \{z = x + iy \in D(0, R) : y > 0\}$. Obviously G_R is a piecewise smooth star shaped region and

$$z_R(t) = \begin{cases} -R + 4Rt & \text{for } 0 \le t \le \frac{1}{2} \\ Re^{i2\pi(t-\frac{1}{2})} & \text{for } \frac{1}{2} \le t \le 1 \end{cases}$$

is a piecewise smooth parameterization of its boundary for which every point in G_R has winding number 1. Furthermore, just as above, by Theorem 6.3.1,

$$\int_0^1 \frac{p\big(z_R(t)\big)}{q\big(z_R(t)\big)}\, dt = i2\pi \sum_{k \in K_+} \frac{p(b_k)}{q'(b_k)}.$$

Hence, since

$$\int_0^{\frac{1}{2}} \frac{p\big(z_R(t)\big)}{q\big(z_R(t)\big)} z_R'(t)\, dt = \int_{-R}^R \frac{p(x)}{q(x)}\, dx \longrightarrow \int_{\mathbb{R}} \frac{p(x)}{q(x)}\, dx \quad \text{as } R \to \infty$$

and

$$\int_{\frac{1}{2}}^1 \frac{p\big(z_R(t)\big)}{q\big(z_R(t)\big)} z_R'(t)\, dt = iR \int_0^\pi \frac{p\big(Re^{i\varphi}\big)}{q\big(Re^{i\varphi}\big)} e^{i\varphi}\, d\varphi \longrightarrow 0 \quad \text{as } R \to \infty,$$

$$\int_{\mathbb{R}} \frac{p(x)}{q(x)}\, dx = \lim_{R\to\infty} \int_0^1 \frac{p\big(z_R(t)\big)}{q\big(z_R(t)\big)} z_R'(t)\, dt = 2\pi \sum_{k \in K_+} \frac{p(b_k)}{q'(b_k)},$$

which is the first equality in the second assertion. Finally, the second equality follows from this and the first assertion. □

To give another example that shows how useful Theorem 6.3.1 is, consider the problem of computing

$$\lim_{R\to\infty} \int_0^R \frac{\sin x}{x}\, dx,$$

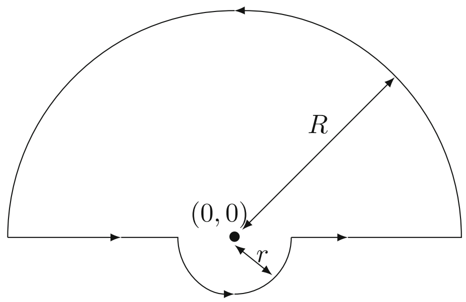
where $\frac{\sin x}{x} \equiv 1$ when $x = 0$. To see that this limit exists, it suffices to show that $\lim_{n\to\infty} \int_0^{n\pi} \frac{\sin x}{x}\, dx$ exists. But if $a_m = (-1)^{m+1} \int_{(m-1)\pi}^{m\pi} \frac{\sin x}{x}\, dx$, then $a_m \searrow 0$ and $\int_0^{n\pi} \frac{\sin x}{x}\, dx = \sum_{m=1}^n (-1)^{m+1} a_m$. Thus Lemma 1.2.1 implies that the limit exists.

We turn now to the problem of evaluation. The first step is to observe that, since $\frac{\sin(-x)}{-x} = \frac{\sin x}{x}$ and $\frac{\cos(-x)}{-x} = -\frac{\cos x}{x}$,

$$2i \int_0^R \frac{\sin x}{x}\, dx = \lim_{r \searrow 0} \left(\int_r^R \frac{e^{ix}}{x}\, dx + \int_{-R}^{-r} \frac{e^{ix}}{x}\, dx \right).$$

Second, for $0 < r < R$, let $G_{r,R}$ be the open set

$$\{z = x + iy \in D(0, R) : \; y > 0\} \cup \{z = x + iy \in D(0, r) : \; y \le 0\} :$$



contour for $\frac{\sin x}{x}$

It should be obvious that $G_{r,R}$ is a piecewise smooth star shaped region. Furthermore, by (6.3.1), 1 is the residue of $\frac{e^{iz}}{z}$ at 0. Now define

$$z_{r,R}(t) = \begin{cases} t & \text{for } -R \le t < -r \\ re^{i(t+r+\pi)} & \text{for } -r \le t < -r + \pi \\ t + 2r - \pi & \text{for } -r + \pi \le t < R - 2r + \pi \\ Re^{i(t-R+2r-\pi)} & \text{for } R - 2r + \pi \le t \le R - 2r + 2\pi. \end{cases}$$

Then the winding number of $t \rightsquigarrow z_{r,R}(t)$ around each $z \in G_{r,R}$ is 1, and so

$$\begin{aligned} i2\pi &= \int_{-R}^{R-2r+2\pi} \frac{e^{iz_{r,R}(t)}}{z_{r,R}(t)} z'_{r,R}(t)\, dt \\ &= \int_{-R}^{-r} \frac{e^{it}}{t}\, dt + i\int_{\pi}^{2\pi} e^{ire^{it}}\, dt + \int_{r}^{R} \frac{e^{it}}{t}\, dt + i\int_{0}^{\pi} e^{iRe^{it}}\, dt \\ &= i2\int_{r}^{R} \frac{\sin t}{t}\, dt + i\int_{\pi}^{2\pi} e^{ire^{it}}\, dt + i\int_{0}^{\pi} e^{iRe^{it}}\, dt. \end{aligned}$$

After letting $r \searrow 0$, we get

$$2\int_0^R \frac{\sin t}{t}\, dt = \pi - \int_0^{\pi} e^{iRe^{it}}\, dt.$$

Finally, note that $\left|e^{iRe^{it}}\right| = e^{-R\sin t}$, and therefore, for any $0 < \delta < \frac{\pi}{2}$,

$$\left|\int_0^\pi e^{iRe^{it}}\,dt\right| \le \int_0^\pi e^{-R\sin t}\,dt = 2\int_0^{\frac{\pi}{2}} e^{-R\sin t}\,dt$$
$$\le 2\delta + 2\int_\delta^{\frac{\pi}{2}} e^{-R\sin t}\,dt \le 2\delta + \pi e^{-R\sin\delta}.$$

Hence, by first letting $R \to \infty$ and then $\delta \searrow 0$, we arrive at

$$\lim_{R\to\infty}\int_0^R \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$$

6.4 The Prime Number Theorem

If p is a prime number (in this discussion, a prime number is an integer greater than or equal to 2 that has no divisors other than itself and 1), then no prime smaller than or equal to p can divide $1 + p!$, and therefore there must exist a prime number that is larger than p. This simple argument, usually credited to Euclid, shows that there are infinitely many prime numbers, but it gives essentially no quantitative information. That is, if $\pi(n)$ is the number of prime numbers $p \le n$, Euclid's argument tells one nothing about how $\pi(n)$ grows as $n \to \infty$. One of the triumphs of late nineteenth century mathematics was the more or less simultaneous proof[1] by Hadamard and de la Vallée Poussin that $\pi(n) \sim \frac{n}{\log n}$ in the sense that

$$\lim_{n\to\infty}\frac{\pi(n)\log n}{n} = 1, \tag{6.4.1}$$

a result that is known as *The Prime Number Theorem*. Hadamard and de Vallée Poussin's proofs were quite involved, and although their proofs were subsequently simplified and other proofs were found, it was not until D.J. Newman came up with the argument given here that a proof of The Prime Number Theorem became accessible to a wide audience. Newman's proof was further simplified by D. Zagier,[2] and it is his version that is presented here.

The strategy of the proof is to show that if $\theta(x) = \sum_{p\le x}\log p$, where the summation is over prime numbers $p \le x$, then

$$\lim_{n\to\infty}\frac{\theta(n)}{n} = 1. \tag{6.4.2}$$

Once (6.4.2) is proved, (6.4.1) is easy. Indeed,

[1]Using empirical evidence, this result had been conjectured by Gauss.

[2]Newman's proof appeared in "Simple Analytic Proof of the Prime Number Theorem", Amer. Math. Monthly, **87** in 1980, and Zagier's paper "Newman's Short Proof of the Prime Number Theorem" appeared in volume **104** of the same journal in 1997, the centennial year of the original proof.

$$\theta(n) = \sum_{p \le n} \log p \le \pi(n) \log n,$$

and so $\frac{\pi(n)\log n}{n} \ge \frac{\theta(n)}{n}$, which, by (6.4.2), means that

$$\underline{\lim}_{n\to\infty} \frac{\pi(n)\log n}{n} \ge 1.$$

At the same time, for each $\alpha \in (0, 1)$,

$$\theta(n) \ge \sum_{n^\alpha < p \le n} \log p \ge \big(\pi(n) - \pi(n^\alpha)\big)\alpha \log n \ge \alpha\big(\pi(n) - n^\alpha\big) \log n,$$

and so, again by (6.4.2),

$$1 = \lim_{n\to\infty} \frac{\theta(n)}{n} \ge \alpha\, \overline{\lim}_{n\to\infty} \frac{\pi(n)\log n}{n}.$$

Since this is true for any $\alpha \in (0, 1)$, (6.4.1) follows.

The first step in proving (6.4.2) is to show that $\theta(x) \le Cx$ for some $C < \infty$ and all $x \ge 0$.

Lemma 6.4.1 *For each $K > \log 2$ there exists an $x_K \in [1, \infty)$ such that $\theta(x) \le 2Kx + \theta(x_K)$ for all $x \ge x_K$. In particular, there exists a $C < \infty$ such that $\theta(x) \le Cx$ for all $x \ge 0$.*

Proof For any $n \in \mathbb{N}$,

$$2^{2n} = (1+1)^{2n} = \sum_{m=0}^{2n} \binom{2n}{m} \ge \binom{2n}{n} = \frac{2^n \prod_{k=1}^n (2k-1)}{n!}.$$

Now set

$$Q = \prod_{n < p \le 2n} p \quad \text{and} \quad D = \frac{2^n \prod_{k=1}^n (2k-1)}{Q}.$$

Because every prime less than or equal to $2n$ divides $2^n \prod_{k=1}^n (2k-1)$, $D \in \mathbb{Z}^+$. Furthermore, because $\frac{Q\times D}{n!} = \binom{2n}{n} \in \mathbb{Z}^+$ and $n!$ is relatively prime to Q, $n!$ must divide D. Hence

$$2^{2n} \ge \frac{Q \times D}{n!} \ge \prod_{n<p\le 2n} p = e^{\theta(2n)-\theta(n)},$$

and so $\theta(2n) - \theta(n) \le 2n \log 2$. Now let $x \ge 2$, and choose $n \in \mathbb{Z}^+$ so that $n - 1 \le \frac{x}{2} \le n$. Then

$$\theta(x) - \theta\left(\tfrac{x}{2}\right) \le \theta(2n) - \theta(n) + \log n \le 2n \log 2 + \log n \le x \log 2 + \log(2x+4).$$

Therefore, for any $K > \log 2$, there exists an $x_K \geq 1$ such that

$$\theta(x) - \theta\left(\frac{x}{2}\right) \leq Kx \quad \text{for } x \geq x_K.$$

Now, for a given $x \geq x_K$, choose $M \in \mathbb{N}$ so that $2^{-M}x \geq x_K > 2^{-M-1}x$. Then

$$\begin{aligned}\theta(x) - \theta(x_K) &\leq \theta(x) - \theta(2^{-M-1}x)\\ &= \sum_{m=0}^{M}\bigl(\theta(2^{-m}x) - \theta(2^{-m-1}x)\bigr) \leq Kx\sum_{m=0}^{M} 2^{-m} \leq 2Kx.\end{aligned}$$

Finally, since $\theta(x) = 0$ for $0 \leq x < 2$, it is clear from this that there exists a $C < \infty$ such that $\theta(x) \leq Cx$ for all $x \geq 0$. □

Lemma 6.4.1 already shows that $\overline{\lim}_{n\to\infty} \frac{\theta(n)}{n} \leq 2\log 2$, and therefore, by the argument we gave to prove the upper bound in (6.4.1) from (6.4.2), $\overline{\lim}_{n\to\infty} \frac{\pi(n)\log n}{n} \leq 2\log 2$. More important, it shows that $\frac{\theta(x)}{x}$ is bounded, and therefore there is a chance that

$$\int_{[1,\infty)} \frac{\theta(x) - x}{x^2}\,dx = \lim_{X\to\infty}\int_1^X \frac{\theta(x) - x}{x^2}\,dx \quad \text{exists in } \mathbb{R}. \tag{6.4.3}$$

Lemma 6.4.2 *If (6.4.3) holds, then (6.4.2) holds.*

Proof Suppose that $\overline{\lim}_{x\to\infty} \frac{\theta(x)}{x} > \lambda$ for some $\lambda > 1$. Then there would exist $\{x_k : k \geq 1\} \subseteq [1,\infty)$ such that $x_k \nearrow \infty$ and $\theta(x_k) > \lambda x_k$. Because θ is non-decreasing, this would mean that

$$\int_{x_k}^{\lambda x_k} \frac{\theta(t) - t}{t^2}\,dt \geq \int_{x_k}^{\lambda x_k} \frac{\lambda x_k - t}{t^2}\,dt = \int_1^{\lambda} \frac{\lambda - t}{t^2}\,dt > 0,$$

and, since

$$\int_{x_k}^{\lambda x_k} \frac{\theta(t) - t}{t^2}\,dt = \int_1^{\lambda x_k} \frac{\theta(t) - t}{t^2}\,dt - \int_1^{x_k} \frac{\theta(t) - t}{t^2}\,dt,$$

this would contradict the existence of the limit in (6.4.3). Similarly, if $\underline{\lim}_{x\to\infty} \frac{\theta(x)}{x} < \lambda$ for some $\lambda < 1$, there would exist $\{x_k : k \geq 1\} \subseteq [1,\infty)$ such that $x_k \nearrow \infty$ and $\theta(x_k) < \lambda x_k$, which would lead to the contradiction that

$$\int_1^{x_k} \frac{\theta(t) - t}{t^2}\,dt - \int_1^{\lambda x_k} \frac{\theta(t) - t}{t^2}\,dt \leq \int_{\lambda}^{1} \frac{\lambda - t}{t^2}\,dt < 0.$$ □

In view of Lemma 6.4.2, everything comes down to verifying (6.4.3). For this purpose, use the change of variables $x = e^t$ to see that

$$\int_1^X \frac{\theta(x)-x}{x^2}\,dx = \int_0^{\log X} \big(e^{-t}\theta(e^t)-1\big)\,dt.$$

Thus, what we have to show is that

$$\int_{[0,\infty)} \big(e^{-t}\theta(e^t)-1\big)\,dt = \lim_{T\to\infty}\int_0^T \big(e^{-t}\theta(e^t)-1\big)\,dt \quad \text{exists in } \mathbb{R}. \tag{6.4.4}$$

Our proof of (6.4.4) relies on a result that allows us to draw conclusions about the behavior of integrals like $\int_0^T \psi(t)\,dt$ as $T\to\infty$ from the behavior of $\int_0^\infty e^{-zt}\psi(t)\,dt$ as $z\to 0$. As distinguished from results that go in the opposite direction, of which (1.10.1) and part (**iv**) of Exercise 3.16 are examples, such results are hard, and, because A. Tauber proved one of the earliest of them, they are known as *Tauberian theorems*. The Tauberian theorem that we will use is the following.

Theorem 6.4.3 *Suppose that* $\psi : [0,\infty) \longrightarrow \mathbb{C}$ *is a bounded function that is Riemann integrable on* $[0,T]$ *for each* $T>0$, *and set*

$$g(z) = \int_{[0,\infty)} e^{-zt}\psi(t)\,dt \quad \textit{for } z\in\mathbb{C} \textit{ with } \Re(z)>0.$$

Then $g(z)$ *is analytic on* $\{z\in\mathbb{C} : \Re(z)>0\}$. *Moreover, if* g *has an extension as an analytic function on an open set containing* $\{z\in\mathbb{C} : \Re(z)\ge 0\}$, *then*

$$\int_{[0,\infty)} \psi(t)\,dt = \lim_{T\to\infty}\int_0^T \psi(t)\,dt$$

exists and is equal to $g(0)$.

Proof For $T>0$, set $g_T(z)=\int_0^T e^{-zt}\psi(t)\,dt$. Using Theorems 3.1.4 and 6.2.6, it is easy to check first that g_T is analytic on $\mathbb{C}$ and then that g is analytic on $\{z\in\mathbb{C} : \Re(z)>0\}$. Thus, what we have to show is that $\lim_{T\to\infty} g_T(0)=g(0)$ if g extends as an analytic function to an open set containing $\{z\in\mathbb{C} : \Re(z)\ge 0\}$. To this end, for $R>0$, choose $\beta_R\in\big(0,\frac{\pi}{2}\big)$ so that g is analytic on an open set containing $\bar{G}$, where

$$G=\{z\in\mathbb{C} : |z|<R \ \&\ \Re(z)>-R\sin\beta_R\}.$$

Clearly (cf. *fig 1* below) G is a piecewise smooth star shaped region, and so we can apply (6.2.5) to the function

$$f(z)\equiv e^{Tz}\left(1+\frac{z^2}{R^2}\right)\big(g(z)-g_T(z)\big)$$

to see that

$$g(0)-g_T(0)=f(0)=\frac{1}{i2\pi}\int_a^b \frac{f\big(z(t)\big)}{z(t)}z'(t)\,dt,$$

where $t \in [a, b] \longmapsto z(t) \in \partial G$ is a piecewise smooth parameterization of ∂G. Furthermore, the parameterization can be chosen so that

$$\int_a^b \frac{f\big(z(t)\big)}{z(t)} z'(t)\, dt = i(J_+ + J_-)$$

where

$$J_+ = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f\big(Re^{i\varphi}\big)\, d\varphi$$

and

$$J_- = \int_{\frac{\pi}{2}}^{\frac{\pi}{2}+\beta_R} f\big(Re^{i\varphi}\big)\, d\varphi - \int_{-R\cos\beta_R}^{R\cos\beta_R} \frac{f\big(-R\sin\beta_R + iy\big)}{-R\sin\beta_R + iy}\, dy + \int_{\frac{3\pi}{2}-\beta_R}^{\frac{3\pi}{2}} f\big(Re^{i\varphi}\big)\, d\varphi.$$

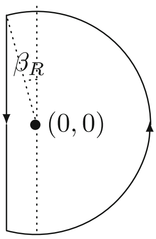


fig 1

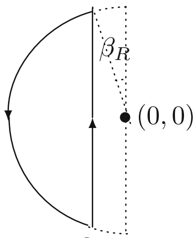


fig 2

To estimate the size of J_+ and J_-, begin by observing that

$$(*) \qquad |z| = R \implies \left|1 + \frac{z^2}{R^2}\right| = \frac{2|\mathfrak{R}(z)|}{R}.$$

To check this, write $z = x + iy$ and conclude that

$$\left|1 + \frac{z^2}{R^2}\right| = \left|\frac{x^2 + y^2 + x^2 - y^2 + i2xy}{R^2}\right| = \left|\frac{2xz}{R^2}\right| = \frac{2|x|}{R}.$$

Now suppose that $|z| = R$ and $\mathfrak{R}(z) > 0$. Then

$$|g(z) - g_T(z)| = \left|\int_{[T,\infty)} e^{-zt}\psi(t)\, dt\right| \le M \int_{[T,\infty)} e^{-\mathfrak{R}(z)t}\, dt \le \frac{Me^{-\mathfrak{R}(z)T}}{\mathfrak{R}(z)},$$

where $M = \|\psi\|_{[0,\infty)}$. Hence, since

$$|f(z)| = e^{\Re(z)T} \left|\left(1 + \frac{z^2}{R^2}\right)\right| |g(z) - g_T(z)|,$$

$|f(Re^{i\varphi})| \leq \frac{2M}{R}$ for $\varphi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and therefore $|J_+| \leq \frac{2\pi M}{R}$. To estimate J_-, we write $f(z) = f_1(z) - f_2(z)$, where

$$f_1(z) = e^{Tz}\left(1 + \frac{z^2}{R^2}\right) g(z) \text{ and } f_2(z) = e^{Tz}\left(1 + \frac{z^2}{R^2}\right) g_T(z),$$

and will estimate the contributions $J_{-,1}$ and $J_{-,2}$ of f_1 and f_2 to J_- separately. First consider the term corresponding to f_2, and remember that g_T is analytic on the whole of $\mathbb{C}$. Thus, by (6.1.3) applied to f_2 on the region (cf. *fig 2* above)

$$\{z \in \mathbb{C} : |z| < R \ \& \ \Re(z) < -R \sin \beta_R\},$$

we know that

$$\int_{-R\cos\beta_R}^{R\cos\beta_R} \frac{f_2\big(-R\sin\beta_R + iy\big)}{-R\sin\beta_R + iy}\, dy + \int_{\frac{\pi}{2}+\beta_R}^{\frac{3\pi}{2}-\beta_R} f_2\big(Re^{i\varphi}\big)\, d\varphi = 0,$$

and therefore that

$$J_{-,2} = \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} f_2\big(Re^{i\varphi}\big)\, d\varphi.$$

Now notice that

$$|g_T(z)| \leq M \int_0^T e^{-\Re(z)t}\, dt = M \frac{e^{-\Re(z)T} - 1}{|\Re(z)|} \leq \frac{Me^{-\Re(z)T}}{|\Re(z)|} \quad \text{if } \Re(z) < 0,$$

and so, by again using $(*)$, we see that $|f_2(Re^{i\varphi})| \leq \frac{2M}{R}$ for $\varphi \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$ and therefore that $|J_{-,2}| \leq \frac{2\pi M}{R}$. Combining this with the estimate for J_+, we now know that

$$|g(0) - g_T(0)| \leq \frac{2M}{R} + \frac{|J_{-,1}|}{2\pi}.$$

Finally, if $B = \|g\|_{\bar{G}}$, then, for $\varphi \in \left[\frac{\pi}{2}, \frac{\pi}{2} + \beta_R\right] \cup \left[\frac{3\pi}{2} - \beta_R, \frac{3\pi}{2}\right]$ and $|y| \leq R\cos\beta_R$,

$$|f_1(Re^{i\varphi})| \leq 4Be^{(\cos\varphi)T} \quad \text{and} \quad \left|\frac{f_1(-R\sin\beta_R + iy)}{-R\sin\beta_R + iy}\right| \leq \frac{4Be^{(-\sin\beta_R)RT}}{R\sin\beta_R},$$

from which it is easy to see that there is a $C_R < \infty$ such that $|J_{-,1}| \leq \frac{C_R}{T}$. Hence, we now have that $|g(0) - g_T(0)| \leq \frac{2M}{R} + \frac{C_R}{2\pi T}$, which means that

$$\varlimsup_{T\to\infty} |g(0) - g_T(0)| \le \frac{2M}{R} \quad \text{for all } R > 0. \qquad \square$$

As the preceding proof demonstrates, not only the choice of contour over which the integral is taken but also the choice of the function f being integrated are crucial for successful applications of Cauchy's integral formula. Indeed, one can replace the f on the right hand side by any analytic function that equals f at the point where one wants to evaluate it. Without such a replacement, the preceding argument would not have worked.

What remains is to show that Theorem 6.4.3 applies when $\psi(t) = e^{-t}\theta(e^t) - 1$. That is, we must show that the function

$$z \rightsquigarrow \int_{[0,\infty)} e^{-zt}\big(e^{-t}\theta(e^t) - 1\big)\,dt$$

on $\{z \in \mathbb{C} : \mathfrak{R}(z) > 0\}$ admits an analytic extension to an open set containing $\{z \in \mathbb{C} : \mathfrak{R}(z) \ge 0\}$. To this end, let $2 = p_1 < \cdots < p_k < \cdots$ be an increasing enumeration of the prime numbers and set $p_0 = 1$. Using summation by parts and taking into account that $\theta(p_0) = 0$ and that $\theta(t) = \theta(p_k)$ for $t \in [p_k, p_{k+1})$, one sees that

$$\begin{aligned}\sum_p \frac{\log p}{p^z} &= \sum_{k=1}^{\infty} \frac{\theta(p_k) - \theta(p_{k-1})}{p_k^z} = \sum_{k=1}^{\infty} \theta(p_k)\left(\frac{1}{p_k^z} - \frac{1}{p_{k+1}^z}\right)\\ &= z\sum_{k=1}^{\infty}\theta(p_k)\int_{p_k}^{p_{k+1}} x^{-z-1}\,dx = z\int_{[2,\infty)} \frac{\theta(x)}{x^{z+1}}\,dx = z\int_{[1,\infty)} \frac{\theta(x)}{x^{z+1}}\,dx\end{aligned}$$

if $\mathfrak{R}(z) > 0$. Hence, after making the change of variables $x = e^t$, we have first that

$$\sum_p \frac{\log p}{p^z} = z\int_{[0,\infty)} e^{-zt}\theta(e^t)\,dt$$

and then that

$$\sum_p \frac{\log p}{p^{z+1}} - \frac{z+1}{z} = (z+1)\int_{[0,\infty)} e^{-zt}\big(e^{-t}\theta(t) - 1\big)\,dt$$

when $\mathfrak{R}(z) > 0$. Now define[3]

$$\Phi(s) = \sum_p \frac{\log p}{p^s} \quad \text{if } \mathfrak{R}(s) > 1,$$

[3] In number theory it is traditional to denote complex numbers by s instead of z, and so we will also.

which, because the series is uniformly convergent on every compact subset of the region $\{s \in \mathbb{C} : \mathfrak{R}(s) > 1\}$, is, by Theorem 6.2.6, analytic. Then the preceding says that

$$(s+1)^{-1}\big(\Phi(s+1) - \tfrac{1}{s} - 1\big) = \int_{[0,\infty)} e^{-st}\big(\theta(t) - 1\big)\,dt,$$

and so checking that Theorem 6.4.3 applies comes down to showing that the function $s \rightsquigarrow \Phi(s) - \frac{1}{s-1}$, which so far is defined only when $\mathfrak{R}(s) > 1$, admits an extension as an analytic function on an open set containing $\{s \in \mathbb{C} : \mathfrak{R}(s) \geq 1\}$.

In order to carry out this program, we will relate the function Φ to the famous function

$$\zeta(s) = \sum_{p} \frac{1}{n^s},$$

which, like Φ, is defined initially only on $\{s \in \mathbb{C} : \mathfrak{R}(s) > 1\}$ and, for the same reason as Φ is, is analytic there. Although this function is usually called the *Riemann zeta function*, it was actually introduced by Euler, who showed that

$$\zeta(s) = \prod_{p} \frac{1}{1 - p^{-s}} \quad \text{if } \mathfrak{R}(s) > 1, \tag{6.4.5}$$

where the product is over all prime numbers. To check (6.4.5), first observe that, when $\mathfrak{R}(s) > 1$, not only the series defining ζ but also the product in (6.4.5) are absolutely convergent. Indeed,

$$\left|\frac{1}{1-p^{-s}} - 1\right| = \frac{p^{-\mathfrak{R}(s)}}{|1 - p^{-s}|} \leq \frac{p^{-\mathfrak{R}(s)}}{1 - 2^{-\mathfrak{R}(s)}},$$

and therefore

$$\sum_{p} \left|\frac{1}{1-p^{-s}} - 1\right| \leq \frac{2^{\mathfrak{R}(s)}}{2^{\mathfrak{R}(s)} - 1} \sum_{n=2}^{\infty} n^{-\mathfrak{R}(s)} < \infty.$$

By Exercises 1.5 and 1.20, this means that both the series and the product are independent of the order in which they are taken. Thus, if we again enumerate the prime numbers in increasing order, $2 = p_1 < \cdots < p_\ell < \cdots$, and take $D(\ell)$ to be the set of $n \in \mathbb{Z}^+$ that are not divisible by any p_k with $k > \ell$, then

$$\zeta(s) = \lim_{\ell\to\infty} \sum_{n \in D_\ell} \frac{1}{n^s} \quad\text{and}\quad \prod_{p} \frac{1}{1-p^{-s}} = \lim_{\ell\to\infty} \prod_{k=1}^{\ell} \frac{1}{1-p_k^{-s}}.$$

The elements n of $D(\ell)$ are in a one-to-one correspondence with $(m_1, \dots, m_\ell) \in \mathbb{N}^\ell$ determined by $n = \prod_{k=1}^{\ell} p_k^{m_k}$. Hence

$$\sum_{n\in D_\ell}\frac{1}{n^s}=\sum_{m_1,\dots,m_\ell=0}^{\infty}\prod_{k=1}^{\ell}p_k^{-sm_k}=\prod_{k=1}^{\ell}\sum_{m_k=0}^{\infty}(p_k^{-s})^{m_k}=\prod_{k=1}^{\ell}\frac{1}{1-p_k^{-s}}.$$

Clearly (6.4.5) follows after one lets $\ell\to\infty$.

Besides establishing a connection between the zeta function and the prime numbers, (6.4.5) shows that, because the product on the right hand side is absolutely convergent and contains no factors that are 0, $\zeta(s)\neq 0$ when $\mathfrak{R}(s)>1$. Both of these properties are important because they allow us to relate ζ to Φ via the equation

$$-\frac{\zeta'(s)}{\zeta(s)}=\sum_p\frac{\log p}{p^s-1}=\Phi(s)+\sum_p\frac{\log p}{p^s(p^s-1)}\quad\text{if }\mathfrak{R}(s)>1.\tag{6.4.6}$$

Since the series $\sum_p\frac{\log p}{p^s(p^s-1)}$ converges uniformly on compact subsets of the region $\{s\in\mathbb{C}:\mathfrak{R}(s)>\frac{1}{2}\}$, it determines an analytic function there. Thus, the extension of $\Phi(s)-\frac{1}{s-1}$ that we are looking for is equivalent to the same sort of extension of $\frac{\zeta'(s)}{\zeta(s)}+\frac{1}{s-1}$, and so an understanding of the zeroes of ζ will obviously play a crucial role.

Lemma 6.4.4 *Set $f(s)=\zeta(s)-\frac{1}{s-1}$ when $\mathfrak{R}(s)>1$. Then f admits an analytic extension to $\{s\in\mathbb{C}:\mathfrak{R}(s)>0\}$, and so ζ admits an analytic extension to the region $\{s\in\mathbb{C}\backslash\{1\}:\mathfrak{R}(s)>0\}$. Furthermore, if $s\neq 1$ and $\mathfrak{R}(s)\geq 1$, then $\zeta(s)\neq 0$. Finally, $\Phi(s)-\frac{1}{s-1}$ extends as an analytic function on an open set containing $\{s\in\mathbb{C}:\mathfrak{R}(s)\geq 1\}$.*

Proof We begin by showing that f admits an analytic extension to $\{s\in\mathbb{C}:\mathfrak{R}(s)>0\}$. To see how this is done, make the change of variables $y=\log x$ to show that if $\mathfrak{R}(s)>1$, then

$$\int_{[1,\infty)}\frac{1}{x^s}\,dx=\int_{[1,\infty)}e^{(1-s)\log x}x^{-1}\,dx=\int_{[0,\infty)}e^{(1-s)y}\,dy=\frac{1}{1-s}.$$

Thus

$$\zeta(s)-\frac{1}{1-s}=\sum_{n=1}^{\infty}\int_n^{n+1}\left(\frac{1}{n^s}-\frac{1}{x^s}\right)dx$$

when $\mathfrak{R}(s)>1$. Since

$$\left|\frac{1}{n^s}-\frac{1}{x^s}\right|=|s|\left|\int_n^x y^{-1-s}\,dy\right|\leq\frac{|s|}{n^{1+\mathfrak{R}(s)}}\quad\text{for }n\leq x\leq n+1,$$

the preceding series converges uniformly on compact subsets of $\{s\in\mathbb{C}:\mathfrak{R}(s)>0\}$ and therefore determines an analytic function there. Thus f admits an analytic extension to $\{s\in\mathbb{C}:\mathfrak{R}(s)>0\}$, and so $s\rightsquigarrow\zeta(s)=f(s)+\frac{1}{s-1}$ admits an analytic extension to $\{s\in\mathbb{C}\backslash\{1\}:\mathfrak{R}(s)>0\}$.

We already know from (6.4.5) that $\zeta(s) \neq 0$ if $\mathfrak{R}(s) > 1$. Now set $z_\alpha = 1 + i\alpha$ for $\alpha \in \mathbb{R}\setminus\{0\}$. Because ζ is analytic in an open disk centered at z_α and is not identically 0 on that disk, (6.2.3) implies that there exists an m_1 which is the smallest $m \in \mathbb{N}$ such that $\zeta^{(m)}(z_\alpha) \neq 0$. Further, using (6.2.3), one sees that

$$m_1 = \lim_{\epsilon \searrow 0} \epsilon \frac{\zeta'(z_\alpha + \epsilon)}{\zeta(z_\alpha + \epsilon)} = -\lim_{\epsilon \searrow 0} \epsilon \Phi(z_\alpha + \epsilon).$$

Because $\zeta(\bar{s}) = \overline{\zeta(s)}$ when $\mathfrak{R}(s) > 1$, $m_1 = -\lim_{\epsilon \searrow 0} \epsilon \Phi(z_{-\alpha} + \epsilon)$. Applying the same argument to $z_{2\alpha}$ and letting m_2 be the smallest $m \in \mathbb{N}$ for which $\zeta^{(m)}(z_{2\alpha}) \neq 0$, we also have that $m_2 = -\lim_{\epsilon \searrow 0} \Phi(z_{\pm 2\alpha} + \epsilon)$. Now remember that the function $f(s) = \zeta(s) - \frac{1}{s-1}$ is analytic on $\{s \in \mathbb{C} : \mathfrak{R}(s) > 0\}$, and therefore

$$\lim_{\epsilon \searrow 0} \epsilon \frac{\zeta'(1+\epsilon)}{\zeta(\epsilon)} = -\lim_{\epsilon \searrow 0} \frac{1 - \epsilon^2 f'(1+\epsilon)}{1 + \epsilon f(1+\epsilon)} = -1,$$

which, by (6.4.6), means that $\lim_{\epsilon \searrow 0} \epsilon \Phi(1+\epsilon) = 1$. Combining these, we have

$$\begin{aligned} -2m_2 - 8m_1 + 6 &= \lim_{\epsilon \searrow 0} \epsilon \sum_{r=-2}^{2} \binom{4}{2+r} \Phi(1 + \epsilon + ir\alpha) \\ &= \lim_{\epsilon \searrow 0} \epsilon \sum_p \frac{(p^{i\alpha} + p^{-i\alpha})^4 \log p}{p^{1+\epsilon}} = 16 \lim_{\epsilon \searrow 0} \sum_p \frac{\big(\cos(\alpha \log p)\big)^4 \log p}{p^{1+\epsilon}} \geq 0, \end{aligned}$$

and this is possible only if $m_1 = 0$. Therefore $\zeta(1 + i\alpha) \neq 0$ for any real $\alpha \neq 0$.

In view of the preceding, we know that, for each $r \in (0, 1)$, $\frac{\zeta'(s)}{\zeta(s)}$ is analytic in an open set containing $\{s \in \mathbb{C}\setminus D(1, r) : \mathfrak{R}(s) \geq 1\}$, and therefore, by (6.4.6), the same is true of $\Phi(s)$. Thus, to show that $\Phi(s) - \frac{1}{s-1}$ extends as an analytic function on an open set containing $\{s \in \mathbb{C} : \mathfrak{R}(s) \geq 1\}$, it suffices for us to show that there is an $r \in (0, 1)$ such that it extends as an analytic function on $D(1, r)$, and this comes down to showing that $\frac{\zeta'(s)}{\zeta(s)} + \frac{1}{s-1}$ admits such an extension. But $\zeta'(s) = f'(s) - (s-1)^{-2}$, and so

$$\frac{\zeta'(s)}{\zeta(s)} + \frac{1}{s-1} = \frac{f(s) + (s-1)f'(s)}{1 + (s-1)f(s)}$$

for s close to 1. Since this means that $\frac{\zeta'(s)}{\zeta(s)} + \frac{1}{s-1}$ stays bounded as $s \to 1$, Theorem 6.2.5 implies that it has an analytic extension to $D(1, r)$ for some $r \in (0, 1)$. □

Lemma 6.4.4 provides the information that was needed in order to apply Theorem 6.4.3, and we have therefore completed the proof of (6.4.1).

There are a couple of comments that should be made about the preceding. The first is the essential role that analytic function theory played. Besides the use of Cauchy's formula in the proof of Theorem 6.4.3, the extension of ζ to $\{s \in \mathbb{C}\setminus\{1\} : \mathfrak{R}(s) > 0\}$

would not have possible if we had restricted our attention to $\mathbb{R}$. Indeed, just staring at the expression $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ for $s \in (1, \infty)$, one would never guess that, as it was in Lemma 6.4.4, sense can be made out of $\zeta\left(\frac{1}{2}\right)$. In the theory of analytic functions, such an extension is called a *meromorphic extension*, and the ability to make them is one of the great benefits of complex analysis. A second comment is about possible refinements of (6.4.1). As our derivation shows, the key to unlocking information about the behavior of $\pi(n)$ for large n is control over the zeroes of the zeta function. Newman's argument allowed us to get (6.4.1) from the fact that $\zeta(s) \neq 0$ if $s \neq 1$ and $\Re(s) \geq 1$. However, in order to get a more refined result, it is necessary to know more about the zeroes of ζ. In this direction, the holy grail is the *Riemann hypothesis*, which is the conjecture that, apart from harmless zeroes, all the zeroes of ζ lie on the vertical line $\{s \in \mathbb{C} : \Re(s) = \frac{1}{2}\}$, but, like any holy grail worthy of that designation, this one remains elusive.

6.5 Exercises

Exercise 6.1 Let G be a connected, open subset of $\mathbb{C}$ that contains a line segment $L = \{\alpha + te^{i\beta} : t \in [-1, 1]\}$ for some $\alpha \in \mathbb{C}$ and $\beta \in \mathbb{R}$. If f and g are analytic functions on G that are equal on L, show that $f = g$ on G.

Exercise 6.2 According to Exercise 5.2,

$$\int_{\mathbb{R}} e^{zx} e^{-\frac{x^2}{2}}\, dx = \sqrt{2\pi} e^{\frac{z^2}{2}}$$

for $z \in \mathbb{R}$. Using Exercise 6.1, show that this equation continues to hold for all $z \in \mathbb{C}$.

Exercise 6.3 Let $f : \mathbb{R} \longrightarrow \mathbb{C}$ be a function that is Riemann integrable on compact intervals and for which

$$\int_{(-\infty,\infty)} f(t)\, dt = \lim_{r \to \infty} \int_{-r}^{r} f(t)\, dt$$

exists in $\mathbb{C}$. Then, by 5.1.2, we know that

$$\int_{\mathbb{R}} f(t + \xi)\, dt = \lim_{r \to \infty} \int_{-r}^{r} f(t + \xi)\, dt = \int_{\mathbb{R}} f(t)\, dt$$

for $\xi \in \mathbb{R}$. The purpose of this exercise is to show that, under suitable conditions, the same equation holds when ξ is replaced by a complex number.

Suppose that f is a continuous function on $\{z \in \mathbb{C} : 0 \leq \Im(z) \leq \eta\}$ that is analytic on $\{z \in \mathbb{C} : 0 < \Im(z) < \eta\}$ for some $\eta > 0$. Further, assume that $\int_{\mathbb{R}} f(t)\, dt \in \mathbb{C}$ exists and that

$$\lim_{r\to\infty}\int_0^\eta f(\pm r+iy)\,dy=0.$$

Using (6.1.3) for the region $\{z=x+iy\in\mathbb{C}:\ |x|<r\ \&\ y\in(0,\eta)\}$, show that

$$\int_{\mathbb{R}} f(t+\xi+i\eta)\,dt=\lim_{r\to\infty}\int_{-r}^{r} f(t+\xi+i\eta)\,dt=\int_{\mathbb{R}} f(t)\,dt$$

for all $\xi\in\mathbb{R}$. Finally, use this to give another derivation of the result in Exercise 6.2.

Exercise 6.4 Here is an elementary algebraic proof of (6.3.2). Let q and $b_1,\dots,b_n$ be as they are there, and use the product representation of q to see that

$$\sum_{k=1}^{n}\frac{q(z)}{q'(b_k)(z-b_k)}-1\quad\text{for } z\in\mathbb{C}\setminus\{b_1,\dots,b_n\}$$

has a unique extension as an at most $(n-1)$st order polynomial P on $\mathbb{C}$. Further, show that $P(b_k)=0$ for $1\le k\le n$, and conclude that P must be 0 everywhere.

Exercise 6.5 Show that

$$\int_0^\pi \frac{1}{a+\cos\varphi}\,d\varphi=\frac{\pi}{\sqrt{a^2-1}}\quad\text{if } a>1.$$

To do this, first note that the integral is half the one when the upper limit is 2π instead of π. Next write the integrand as $\frac{2e^{i\varphi}}{2ae^{i\varphi}+e^{i2\varphi}+1}$, and conclude that

$$\int_0^\pi \frac{1}{a+\cos\varphi}\,d\varphi=\int_0^{2\pi}\frac{e^{i\varphi}}{e^{i2\varphi}+2ae^{i\varphi}+1}\,d\varphi.$$

Now apply (6.3.2) and (6.2.5) to complete the calculation.

Exercise 6.6 Show that

$$\int_0^{2\pi}\frac{1}{a^2+\sin^2\varphi}\,d\varphi=\pi\left(\frac{1}{\sqrt{a^2+1}}+\frac{1}{\sqrt{a^2-1}}\right)\quad\text{if } |a|>1.$$

This calculation can be reduced to one like that in Exercise 6.5 by first noting that $a^2+\sin^2\varphi=(a+i\sin\varphi)(a-i\sin\varphi)$.

Exercise 6.7 Show that

$$\int_{\mathbb{R}}\frac{1}{1+x^4}\,dx=\frac{\pi}{\sqrt{2}},$$

that

$$\int_{\mathbb{R}}\frac{\cos(\alpha x)}{(1+x^2)^2}\,dx=\frac{\pi(1+\alpha)}{2e^\alpha}\quad\text{for } \alpha\ge 0$$

and that

$$\int_{\mathbb{R}} \frac{\sin^2 x}{(1+x^2)^2}\,dx = \frac{\pi(e^2-3)}{4e^2}.$$

Exercise 6.8 Let $\mu \in (0,1)$ and $a \in \mathbb{C}\setminus\mathbb{Z}$, and define

$$f(z) = \frac{e^{i\pi(2\mu-1)z}}{(z-a)\sin(\pi z)} \quad \text{for } z \in \mathbb{C}\setminus(\mathbb{Z}\cup\{a\}).$$

By applying Theorem 6.3.1 to the integral of f around the circle $\mathbb{S}^1\big(0, n+\frac{1}{2}\big)$ for $n > a$ and then letting $n \to \infty$, show that

$$\frac{1}{\pi}\sum_{k=-\infty}^{\infty} \frac{e^{i2\pi\mu k}}{a-k} = \frac{e^{i\pi(2\mu-1)a}}{\sin(\pi a)}.$$

Exercise 6.9 Let $z : [a,b] \longrightarrow \mathbb{C}$ be a closed, piecewise smooth path, and set

$$N(z) = \frac{1}{i2\pi}\int_a^b \frac{z'(t)}{z(t)-z}\,dt$$

for $z \notin \{z(t) : t \in [a,b]\}$. As we saw in the discussion following Lemma 6.2.9, $N(z)$ is the number of times that the path winds around z, but there is another way of seeing that $N(z)$ must be an integer. Namely, set

$$\alpha(t) = \int_a^t \frac{z'(\tau)}{z(\tau)-z}\,d\tau,$$

and consider the function $u(t) = e^{-\alpha(t)}\big(z(t)-z\big)$. Show that $u'(t) = 0$ for all but a finite number of $t \in [a,b]$, and conclude that $e^{\alpha(t)} = \frac{z(t)-z}{z(a)-z}$. In particular, $e^{\alpha(b)} = 1$, and so $\alpha(b)$ must be an integer multiple of $i2\pi$. Next show that if G is a connected open subset of $\mathbb{C}\setminus\{z(t) : t \in [a,b]\}$, then $z \rightsquigarrow N(z)$ is constant on G.

Exercise 6.10 Let $\mathcal{H}$ be the space of analytic functions f on $\mathbb{C}$ for which

$$\|f\|_{\mathcal{H}} \equiv \sqrt{\frac{1}{\pi}\int_{\mathbb{C}} |f(z)|^2 e^{-|z|^2}\,dx\,dy} < \infty.$$

It should be clear that $\mathcal{H}$ is a vector space over $\mathbb{C}$ (i.e., linear combinations of its elements are again elements). In addition, it has an inner product given by

$$(f,g)_{\mathcal{H}} \equiv \frac{1}{\pi}\int_{\mathbb{C}} f(z)\overline{g(z)}e^{-|z|^2}\,dx\,dy.$$

(**i**) For $r > 0$ and $\zeta \in \mathbb{C}$, set

$$M_r(\zeta) = \frac{1}{\pi r^2} \int_{\overline{D(\zeta,r)}} f(z)\, dx\, dy,$$

show that

$$0 \le \frac{1}{\pi r^2} \int_{\overline{D(\zeta,r)}} |f(z) - M_r(\zeta)|^2\, dx\, dy = \frac{1}{\pi r^2} \int_{\overline{D(\zeta,r)}} |f(z)|^2\, dx\, dy - |M_r(\zeta)|^2,$$

and, using (6.2.2), conclude that

$$\sup_{\zeta \in \overline{D(0,r)}} |f(\zeta)| \le \frac{1}{4\pi r^2} \int_{\overline{D(0,2r)}} |f(z)|\, dx\, dy \le \frac{e^{4r^2} \|f\|_{\mathcal{H}}}{4\pi r^2}. \tag{6.5.1}$$

(**ii**) From (6.5.1) we know that $\|f\|_{\mathcal{H}} = 0$ if and only if $f = 0$. Further, using the same reasoning as was used in Sect. 4.1 in connection with Schwarz's inequality and the triangle inequality, show that

$$|(f, g)_{\mathcal{H}}| \le \|f\|_{\mathcal{H}} \|g\|_{\mathcal{H}} \text{ and } \|\alpha f + \beta g\|_{\mathcal{H}} \le |\alpha| \|f\|_{\mathcal{H}} + |\beta| \|g\|_{\mathcal{H}}.$$

(**iii**) Let $\{f_n : n \ge 1\}$ be sequence in $\mathcal{H}$, and say that $\{f_n : n \ge 1\}$ converges in $\mathcal{H}$ to $f \in \mathcal{H}$ if $\lim_{n\to\infty} \|f_n - f\|_{\mathcal{H}} = 0$. Show that $\mathcal{H}$ with this notion of convergence is complete. That is, show that if

$$\lim_{m\to\infty} \sup_{n>m} \|f_n - f_m\|_{\mathcal{H}} = 0,$$

then there exists an $f \in \mathcal{H}$ to which $\{f_n : n \ge 1\}$ converges in $\mathcal{H}$. Here are some steps that you might want to take. First, using (6.5.1) applied to $f_n - f_m$, the obvious analog of Lemma 1.4.4 for functions on $\mathbb{C}$ combined with Theorem 6.2.6 show that there is an analytic function f on $\mathbb{C}$ to which $\{f_n : n \ge 1\}$ converges uniformly on compact subsets. Next, given $\epsilon > 0$, choose m_ϵ so that $\|f_n - f_m\|_{\mathcal{H}} < \epsilon$ for $n > m \ge m_\epsilon$, and then show that

$$\left(\int_{\overline{D(0,r)}} |f(z) - f_m(z)|^2 e^{-|z|^2}\, dx\, dy \right)^{\frac{1}{2}} = \lim_{n\to\infty} \left(\int_{\overline{D(0,r)}} |f_n(z) - f_m(z)|^2 e^{-|z|^2}\, dx dy \right)^{\frac{1}{2}} \le \epsilon$$

for all $m \ge m_\epsilon$ and $r > 0$. Finally, conclude that $f \in \mathcal{H}$ and $\|f - f_m\|_{\mathcal{H}} \le \epsilon$ for all $m \ge m_\epsilon$.

A vector space that has an inner product for which the associated notion of convergence makes it complete is called a *Hilbert space*, and the Hilbert space $\mathcal{H}$ is known as the *Fock space*.

Exercise 6.11 This exercise is a continuation of Exercise 6.10. We already know that $\mathcal{H}$ is a Hilbert space, and the goal here is to produce an orthonormal basis for $\mathcal{H}$.

(**i**) Set $\partial = \frac{1}{2}(\partial_x - i\partial_y)$, note that $\partial^m e^{-|z|^2} = \overline{z^m} e^{-|z|^2}$, and use integration by parts to show that

$$(*) \qquad \int_{\mathbb{C}} f(z)\overline{z^m} e^{-|z|^2}\, dx\, dy = \int_{\mathbb{C}} f^{(m)}(z) e^{-|z|^2}\, dx\, dy \quad \text{for } m \geq 0 \text{ and } f \in \mathcal{H}.$$

(**ii**) Use $(*)$ and 5.4.3 to show that

$$\int_{\mathbb{C}} z^m \overline{z^n} e^{-|z|^2}\, dx\, dy = \delta_{m,n} \pi\, m! \quad \text{for } m,\ n \in \mathbb{N}.$$

Next, set $\mathbf{e}_m(z) = (m!)^{-\frac{1}{2}} z^m$, and, using the fact that $(\mathbf{e}_m, \mathbf{e}_n)_{\mathcal{H}} = \delta_{m,n}$, show that the $\mathbf{e}_m$'s are linearly independent in the sense that for any $n \geq 1$ and $\alpha_0, \dots, \alpha_n \in \mathbb{C}$,

$$\sum_{m=0}^{n} \alpha_m \mathbf{e}_m = 0 \implies \alpha_0 = \cdots = \alpha_n = 0.$$

Use this to prove that $\{\mathbf{e}_n : n \geq 0\}$ is a bounded sequence in $\mathcal{H}$ that admits no convergent subsequence.

(**iii**) The preceding proves that $\mathcal{H}$ is *not* finite dimensional, and the concluding remark there highlights one of the important ways in which infinite dimensional vector spaces are distinguished from finite dimensional ones. At the same time, it opens the possibility that $(\mathbf{e}_0, \dots, \mathbf{e}_m, \dots)$ is an orthonormal basis. That is, we are asking whether, in an appropriate sense,

$$(**) \qquad f = \sum_{m=0}^{\infty} (f, \mathbf{e}_m)_{\mathcal{H}} \mathbf{e}_m \quad \text{for all } f \in \mathcal{H}.$$

To answer this question, use $(*)$, polar coordinates, and (6.2.1) to see that

$$\int_{\mathbb{C}} f(z)\overline{z^m} e^{-|z|^2}\, dx\, dy = \pi f^{(m)}(0)$$

and therefore that

$$\sum_{m=0}^{\infty} (f, \mathbf{e}_m)_{\mathcal{H}} \mathbf{e}_m(z) = \sum_{m=0}^{\infty} \frac{f^{(m)}(0)}{m!} z^m = f(z).$$

Hence, $(**)$ holds in the sense that the series on the right hand side converges uniformly on compact subsets to the function on the left hand side.

(**iv**) Given $f \in \mathcal{H}$, set $f_n = \sum_{m=0}^{n}(f, \mathbf{e}_m)_{\mathcal{H}}\mathbf{e}_m$. In (**iii**) we showed that $f_n \longrightarrow f$ uniformly on compact subsets, and the goal here is to show that $\|f_n - f\|_{\mathcal{H}} \longrightarrow 0$. The strategy is to show that there is a $g \in \mathcal{H}$ such that $\|f_n - g\|_{\mathcal{H}} \longrightarrow 0$. Once one knows that such a g exists, it essentially trivial to show that $g = f$. Indeed, by applying (6.5.1) to $g - f_n$, one knows that $f_n \longrightarrow g$ uniformly on compact subsets and therefore, since $f_n \longrightarrow f$ uniformly on compact subsets, it follows that $f = g$. One way to prove that $\{f_n : n \geq 0\}$ converges in $\mathcal{H}$ is to use Cauchy's criterion (cf. (**iii**) in Exercise 6.10). To this end, show that $(f_n, f - f_n)_{\mathcal{H}} = 0$ for all $n \geq 0$, and use this to show that

$$\|f\|_{\mathcal{H}}^2 = \|f_n\|_{\mathcal{H}}^2 + \|f - f_n\|_{\mathcal{H}}^2 \geq \|f_n\|_{\mathcal{H}}^2 = \sum_{m=0}^{n} |(f, \mathbf{e}_m)_{\mathcal{H}}|^2.$$

From this conclude that $\sum_{m=0}^{\infty} |(f, \mathbf{e}_m)_{\mathcal{H}}|^2 < \infty$ and therefore

$$\sup_{n>m} \|f_n - f_m\|_{\mathcal{H}}^2 = \sup_{n>m} \sum_{k=m+1}^{n} |(f, \mathbf{e}_k)_{\mathcal{H}}|^2 \longrightarrow 0 \quad \text{as } m \to \infty.$$

Appendix

The goal here is to construct, starting from the set $\mathbb{Q}$ of rational numbers, a model for the real line. That is, we want to construct a set $\mathbb{R}$ that comes equipped with the following properties.

(1) There is a one-to-one embedding Φ taking $\mathbb{Q}$ into $\mathbb{R}$.
(2) There is an order relation "$<$" on $\mathbb{R}$ such that, for all $s, t \in \mathbb{Q}$, $\Phi(r) < \Phi(s)$ if and only if $s < r$ and, for all $x, y \in \mathbb{R}$, $x < y$ or $x = y$ or $y < x$ (i.e., $x > y$).
(3) There are arithmetic operations $(x, y) \in \mathbb{R}^2 \longmapsto (x + y) \in \mathbb{R}$ and $(x, y) \in \mathbb{R}^2 \longmapsto xy \in \mathbb{R}$ such that $\Phi(r + s) = \Phi(r) + \Phi(s)$ and $\Phi(rs) = \Phi(r)\Phi(s)$ for all $r, s \in \mathbb{Q}$. Furthermore, these operations have the same properties as the corresponding operations on $\mathbb{Q}$.
(4) Define $|x| \equiv \pm x$ depending on whether $x \geq \Phi(0)$ (i.e., $x = \Phi(0)$ or $x > \Phi(0)$) or $x < \Phi(0)$, and say that a sequence $\{x_n : n \geq 1\} \subseteq \mathbb{R}$ converges to $x \in \mathbb{R}$ and write $x_n \longrightarrow x$ if for each $r \in \mathbb{Q}^+ \equiv \{r \in \mathbb{Q} : r > 0\}$ there is an m such that $|x - x_n| < r$ when $n \geq m$. Then for every $x \in \mathbb{R}$ there exists a sequence $\{r_n : n \geq 1\} \subseteq \mathbb{Q}$ such that $\Phi(r_n) \longrightarrow x$.
(5) The space $\mathbb{R}$ endowed with the preceding notion of convergence is complete. That is, if $\{x_n : n \geq 1\} \subseteq \mathbb{R}$ and, for each $r \in \mathbb{Q}^+$ there exists an m such that $|x_n - x_m| < r$ when $n \geq m$, then there exists an $x \in \mathbb{R}$ to which $\{x_n : n \geq 1\}$ converges.

Obviously, if it were not for property (5), there would be no reason not to take $\mathbb{R} = \mathbb{Q}$. Thus, the challenge is to embed $\mathbb{Q}$ in a structure for which Cauchy's criterion guarantees convergence. There are several ways to build such a structure, the most popular being by a method known as "Dedekind cuts". However, we will use a method that is based on ideas of Hausdorff.

Use $\tilde{\mathbb{R}}$ to denote the set of all maps, which should be thought of sequences, $S : \mathbb{Z}^+ \longrightarrow \mathbb{Q}$ with the property that for all $r \in \mathbb{Q}^+$ there exists an $i \in \mathbb{Z}^+$ such that $|S(j) - S(i)| \leq r$ if $j > i$. Given $S, S' \in \tilde{\mathbb{R}}$, say that S' is equivalent to S and write $S' \sim S$ if for each $r \in \mathbb{Q}^+$ there exists an $i \in \mathbb{Z}^+$ such that $|S'(j) - S(j)| \leq r$ for $j \geq i$. Then $\sim$ is an equivalence relation on $\tilde{\mathbb{R}}$ in the sense that, for all $S, S' \in \tilde{\mathbb{R}}$,

D.W. Stroock, *A Concise Introduction to Analysis*,
DOI 10.1007/978-3-319-24469-3

$S \sim S$, $S' \sim S \iff S \sim S'$, and $S' \sim S$ if there exists an $S'' \in \tilde{\mathbb{R}}$ such that $S \sim S''$ and $S' \sim S''$. The first two of these are obvious. To check the third, for a given $r \in \mathbb{Q}^+$ choose i so that $|S(j) - S''(j)| \vee |S'(j) - S''(j)| \le \frac{r}{2}$ for $j \ge i$. Then

$$|S'(j) - S(j)| \le |S'(j) - S''(j)| + |S''(j) - S(j)| \le r \quad \text{for } j \ge i.$$

Lemma A.1 *If $S \in \tilde{\mathbb{R}}$, $\{j_k : k \ge 1\} \subseteq \mathbb{Z}^+$ is a strictly increasing sequence, and $S'(k) = S(j_k)$ for $k \ge 1$, then $S' \in \tilde{\mathbb{R}}$ and $S' \sim S$.*

Proof It is obvious that $S' \in \tilde{\mathbb{R}}$. To see that $S' \sim S$, given $r \in \mathbb{Q}^+$, choose i so that $|S(j) - S(i)| \le \frac{r}{2}$ for $j \ge i$. Then

$$|S(j_k) - S(k)| \le |S(j_k) - S(i)| + |S(k) - S(i)| \le r \quad \text{for } k \ge i. \qquad \square$$

Given a sequence $\{S_n : n \ge 1\} \subseteq \tilde{\mathbb{R}}$, we will say that $\{S_n : n \ge 1\}$ converges to $S \in \tilde{\mathbb{R}}$ and will write $S_n \longrightarrow S$ if, for each $r \in \mathbb{Q}^+$, there exists an $(m, i) \in (\mathbb{Z}^+)^2$ such that $|S_n(j) - S(j)| \le r$ for $n \ge m$ and $j \ge i$. Notice that $\{S_n : n \ge 1\}$ can converge to more than one element of $\tilde{\mathbb{R}}$. However, if it converges to both S and S', then $S' \sim S$. Next, say that $\{S_n : n \ge 1\}$ is Cauchy convergent if for all $r \in \mathbb{Q}^+$ there exists an $(m, i) \in (\mathbb{Z}^+)^2$ such that $|S_n(j) - S_m(j)| \le r$ for all $n \ge m$ and $j \ge i$. It is easy to check that if $\{S_n : n \ge 1\}$ converges to some S, then it is Cauchy convergent.

Lemma A.2 *Assume that $\{S_n : n \ge 1\} \subseteq \tilde{\mathbb{R}}$ is Cauchy convergent. Then for each $r \in \mathbb{Q}^+$ there exists an i such that $|S_n(j) - S_n(i)| \le r$ for all $n \ge 1$ and $j \ge i$. Furthermore, if $S'_n \sim S_n$ for each $n \in \mathbb{Z}^+$ and $\{S'_n : n \ge 1\}$ is Cauchy convergent, then for each $r \in \mathbb{Q}^+$ there exists an $(m, i) \in (\mathbb{Z}^+)^2$ such that $|S'_n(j) - S_n(j)| \le r$ for all $n \ge m$ and $j \ge i$. In particular, if $S_n \longrightarrow S$ and $S'_n \longrightarrow S'$, then $S' \sim S$.*

Proof Given $r \in \mathbb{Q}^+$, choose $(m, i') \in (\mathbb{Z}^+)^2$ so that $|S_n(j) - S_m(j)| \le \frac{r}{3}$ for $n \ge m$ and $j \ge i'$. Next choose $i \ge i'$ so that $|S_\ell(j) - S_\ell(i)| \le \frac{r}{3}$ for $1 \le \ell \le m$ and $j \ge i$. Then, for $n \ge m$ and $j \ge i$,

$$|S_n(j) - S_n(i)| \le |S_n(j) - S_m(j)| + |S_m(j) - S_m(i)| + |S_m(i) - S_n(i)| \le r.$$

Hence $|S_n(j) - S_n(i)| \le r$ for all $n \ge 1$ and $j \ge i$.

Now assume that $S'_n \sim S_n$ for all n and that $\{S'_n : n \ge 1\}$ is Cauchy convergent. Given $r \in \mathbb{Q}^+$, use the preceding to choose i' so that

$$|S'_n(j_2) - S'_n(j_1)| \vee |S_n(j_2) - S_n(j_1)| \le \frac{r}{5} \quad \text{for all } n \ge 1 \text{ and } j_1,\ j_2 \ge i',$$

and then choose m and $i \ge i'$ so that $|S'_m(i) - S_m(i)| \le \frac{r}{5}$ and

$$|S'_n(j) - S'_m(j)| \vee |S_n(j) - S_m(j)| \le \frac{r}{5} \quad \text{for } n \ge m \text{ and } j \ge i.$$

Then, for $n \geq m$ and $j \geq i$,

$$\begin{aligned}|S'_n(j) - S_n(j)| \leq |S'_n(j) - S'_m(j)| + |S'_m(j) - S'_m(i)| + |S'_m(i) - S_m(i)| \\ + |S_m(i) - S_m(j)| + |S_m(j) - S_n(j)| \leq r.\end{aligned}$$

Finally, if $S_n \longrightarrow S$ and $S'_n \longrightarrow S'$, then for any $n \geq m$ and $j \geq i$

$$\begin{aligned}|S'(j) - S(j)| &\leq |S'(j) - S'_n(j)| + |S'_n(j) - S_n(j)| + |S_n(j) - S(j)| \\ &\leq r + |S'(j) - S'_n(j)| + |S(j) - S_n(j)|,\end{aligned}$$

and so, by taking n sufficiently large, we conclude that $|S'(j) - S(j)| \leq 2r$ for all sufficiently large j. □

Lemma A.3 *If* $\{S_n : n \geq 1\} \subseteq \tilde{\mathbb{R}}$ *is Cauchy convergent, then there exists an* $S \in \tilde{\mathbb{R}}$ *to which it converges.*

Proof By Lemma A.2, there exists a sequence $\{i'_k : k \geq 1\} \subseteq \mathbb{Z}^+$ which is strictly increasing for which $|S_n(j) - S_n(i'_k)| \leq \frac{1}{k}$ for all $n \geq 1$ and $j \geq i'_k$. Next, because $\{S_n : n \geq 1\}$ is Cauchy convergent, we can choose strictly increasing sequences $\{m_k : k \geq 1\} \subseteq \mathbb{Z}^+$ and $\{i_k : k \geq 1\}$ such that, for each $k \geq 1$, $i_k \geq i'_k$ and $|S_n(j) - S_{m_k}(j)| \leq \frac{1}{k}$ for all $n \geq m_k$ and $j \geq i_k$. Define $S(j) = S_{m_j}(j)$ for $j \in \mathbb{Z}^+$. It is obvious that $S \in \tilde{\mathbb{R}}$. In addition, for $n \geq m_k$ and $j \geq i_k$,

$$\begin{aligned}|S_n(j) - S(j)| \leq |S_n(j) - S_n(i_k)| + |S_n(i_k) - S_{m_k}(i_k)| \\ + |S_{m_k}(i_k) - S_{m_k}(i_j)| + |S_{m_k}(i_j) - S_{m_j}(i_j)| \leq \tfrac{4}{k}.\end{aligned}$$ □

We are now ready to describe our model for $\mathbb{R}$. Namely, for each $S \in \tilde{\mathbb{R}}$, let $[S] \equiv \{S' \in \tilde{\mathbb{R}} : S' \sim S\}$ be the equivalence class of S, and take $\mathbb{R}$ to be the set $\{[S] : S \in \tilde{\mathbb{R}}\}$ of equivalences classes. Because $\sim$ is an equivalence relation, it is easy to check that $S \in [S]$ and that either $[S] = [S']$ or $[S'] \cap [S] = \emptyset$. Thus $\mathbb{R}$ is a partition of $\tilde{\mathbb{R}}$ into mutually disjoint, non-empty subsets. We next embed $\mathbb{Q}$ into $\mathbb{R}$ by identifying $r \in \mathbb{Q}$ with $[\mathbf{R}]$, where $\mathbf{R}$ is the element of $\tilde{\mathbb{R}}$ such that $\mathbf{R}(j) = r$ for all $j \in \mathbb{Z}^+$. That is, the map Φ in (1) is given by $\Phi(r) = [\mathbf{R}]$. Although it entails an abuse of notation, we will often use r to play a dual role: it will denote an element of $\mathbb{Q}$ as well as the associated element $\Phi(r) = [\mathbf{R}]$ of $\mathbb{R}$.

The next step is to introduce an arithmetic structure on $\mathbb{R}$. To this end, given $S,\ T \in \tilde{\mathbb{R}}$ define $S + T$ and ST to be the elements of $\tilde{\mathbb{R}}$ such that $(S + T)(j) = S(j) + T(j)$ and $(ST)(j) = S(j)T(j)$ for $j \in \mathbb{Z}^+$. It is easy to check that if $S' \sim S$ and $T' \sim T$, then $S' + T' \sim S + T$ and $S'T' \sim ST$. Thus, given $x,\ y \in \mathbb{R}$, where $x = [S]$ and $y = [T]$, we can define $x + y$ and xy unambiguously by $x + y = [S + T]$ and $xy = [ST]$, in which case it is easy to check that $x + y = y + x$, $xy = yx$, $(x + y) + z = x + (y + z)$, and $(x + y)z = xz + yz$. Similarly, if $x = [S]$, then we can define $-x = [-S]$, in which case $x + y = 0$ if and only if $y = -x$. Indeed, it is obvious that $y = -x \implies x + y = 0$. Conversely, if $x = [S]$, $y = [T]$, and

$x+y=0$, then for all $r \in \mathbb{Q}^+$ there exists an i such that $|S(j)+T(j)| \le r$ for $j \ge i$. Since this means that $\left|T(j) - \bigl(-S(j)\bigr)\right| \le r$ for $j \ge i$, it follows that $T \sim -S$ and therefore that $y = [-x]$. Finally, it is easy to check that $\Phi(r+s) = \Phi(r) + \Phi(s)$ and $\Phi(rs) = \Phi(r)\Phi(s)$ for $r,\ s \in \mathbb{Q}$ and that $\Phi(-r) = -\Phi(r)$ for $r \in \mathbb{Q}$.

Lemma A.4 *For any $S \in \tilde{\mathbb{R}}$, $[S] \ne 0$ if and only if there exists an $r \in \mathbb{Q}^+$ and an i such that either $S(j) \le -r$ for all $j \ge i$ or $S(j) \ge r$ for all $j \ge i$. Moreover, if $x,\ y \in \mathbb{R}$, then $xy = 0$ if and only if either $x = 0$ or $y = 0$. Finally, if $x \ne 0$, then there exists a unique $\frac{1}{x} \in \mathbb{R}$ such that $x\frac{1}{x} = 1$, and $\Phi\bigl(\frac{1}{r}\bigr) = \frac{1}{\Phi(r)}$ for $r \in \mathbb{Q} \setminus \{0\}$.*

Proof Because $[S] \ne 0$, there exists an $r \in \mathbb{Q}^+$ and arbitrarily large $j \in \mathbb{Z}^+$ for which $|S(j)| \ge r$. Now choose i so that $|S(j) - S(i)| \le \frac{r}{4}$ for $j \ge i$. If $S(j) \ge r$ for some $j \ge i$, then $S(j) \ge \frac{r}{2}$ for all $j \ge i$. Similarly, if $S(j) \le -r$ for some $j \ge i$, then $S(j) \le -\frac{r}{2}$ for all $j \ge i$.

It is obvious that $x0 = 0$ for any $x \in \mathbb{R}$. Now suppose that $xy = 0$, where $x = [S]$ and $y = [T]$. If $x \ne 0$, use the preceding to choose $r_0 \in \mathbb{Q}^+$ and i_0 such that $|S(j)| \ge r_0$ for $j \ge i_0$. Then $|S(j)T(j)| \ge r_0|T(j)|$ for all $j \ge i_0$. Since $xy = 0$, for any $r \in \mathbb{Q}^+$ there exists an $i \ge i_0$ such that $|S(j)T(j)| \le r_0 r$ and therefore $|T(j)| \le r$ for $j \ge i$. Hence $T \sim \mathbf{0}$ and so $y = 0$.

Finally, assume that $x \ne 0$. To see that there is at most one y for which $xy = 1$, suppose that $xy_1 = 1 = xy_2$. Then $x(y_1 - y_2) = 0$, and so, since $x \ne 0$, $y_1 - y_2 = 0$. But this means that $-y_2 = -y_1$ and therefore that $y_1 = y_2$. To construct a y for which $xy = 1$, suppose that $x = [S]$ and choose $r \in \mathbb{Q}^+$ and i so that $|S(j)| \ge r$ for $j \ge i$. Next, define S' so that $S'(j) = r$ for $1 \le j \le i$ and $S'(j) = S(j)$ for $j > i$. Then $S' \sim S$, and so $x = [S']$. Finally, define $T(j) = \frac{1}{S'(j)}$ for all $j \in \mathbb{Z}^+$, and observe that $T \in \tilde{\mathbb{R}}$ and $S'T = \mathbf{1}$. Hence $x[T] = 1$, and so we can take $\frac{1}{x} = [T]$. Obviously, $\Phi\bigl(\frac{1}{r}\bigr) = \frac{1}{\Phi(r)}$ for $r \in \mathbb{Q} \setminus \{0\}$. □

We now introduce an order relation on $\mathbb{R}$. Given $S,\ T \in \tilde{\mathbb{R}}$, write $S < T$ if there exists an $r \in \mathbb{Q}^+$ and an i such that $S(j) + r \le T(j)$ for $j \ge i$. If $S' \sim S$ and $T' \sim T$, then $S < T \implies S' < T'$. Indeed, choose $r \in \mathbb{Q}^+$ and i_0 so that $S(j) + r \le T(j)$ for all $j \ge i_0$, and then choose $i \ge i_0$ so that $|S'(j) - S(j)| \vee |T'(j) - T(j)| \le \frac{r}{4}$ for $j \ge i$. Then $S'(j) + \frac{r}{2} \le T'(j)$ for $j \ge i$. Thus, we can write $x < y$ if $x = [S]$ and $y = [T]$ with $S < T$. Further, if we write $x > y$ when $y < x$, then we can say that, for any $x,\ y \in \mathbb{R}$, $x < y$, $x > y$, or $x = y$. To see this, assume that $x \ne y$. If $x = [S]$ and $y = [T]$, then, by Lemma A.4, there exists an $r \in \mathbb{Q}^+$ and i such that $S(j) - T(j) \le -r$ for all $j \ge i$ or $S(j) - T(j) \ge r$ for all $j \ge i$. In the first case $x < y$ and in the second $x > y$. It is easy to check that $x < y \iff -x > -y$, and clearly, for any $r,\ s \in \mathbb{Q}$, $r < s \iff \Phi(r) < \Phi(s)$. Finally, we will write $x \le y$ if $x < y$ or $x = y$ and $x \ge y$ if $y \le x$.

Define $|x|$ as in (4). Using Lemma A.4, one sees that if $x = [S]$ then $|x| = [|S|]$. Clearly $x = 0 \iff |x| = 0$. In fact, $x = 0$ if and only if $|x| \le r$ for all $r \in \mathbb{Q}^+$. In addition, $|xy| = |x||y|$ and the triangle inequality $|x+y| \le |x| + |y|$ holds for all $x,\ y \in \mathbb{R}$. The first of these is trivial. To see the second, assume that $|x+y| \ne |x| + |y|$. Then either $|x+y| < |x| + |y|$ or $|x+y| > |x| + |y|$.

But if $|x + y| > |x| + |y|$, $x = [S]$, and $y = [T]$, then we would have the contradiction that $|S(j)+T(j)| > |S(j)|+|T(j)|$ for large enough j. Now introduce the notion of convergence described in (4). To see that $\{x_n : n \ge 1\}$ can converge to at most one x, suppose that it converges to x and x'. Then, by the triangle inequality, $|x' - x| \le |x' - x_n| + |x_n - x|$ for all n, and so $|x' - x| \le r$ for all $r \in \mathbb{Q}^+$. Finally, notice that $|(x' + y) - (x + y)| = |x' - x|$, $|x'y - xy| = |x' - x||y|$, and, if $xx' \ne 0$, $\left|\frac{1}{x'} - \frac{1}{x}\right| = \frac{|x'-x|}{|x'||x|}$, which means that the addition and multiplication operations on $\mathbb{R}$ as well as division on $\mathbb{R} \setminus \{0\}$ are continuous with respect to this notion of convergence.

Lemma A.5 *For every $x \in \mathbb{R}$ there exists a sequence $\{r_n : n \ge 1\} \subseteq \mathbb{Q}$ such that $r_n \longrightarrow x$, and so $\Phi(\mathbb{Q})$ is dense in $\mathbb{R}$.*

Proof Suppose that $x = [S]$, and set $r_n = [\mathbf{R}_n]$ where $\mathbf{R}_n(j) = S(n)$ for all $j \in \mathbb{Z}^+$. Given $r \in \mathbb{Q}^+$, choose m so that $|S(n) - S(m)| \le \frac{r}{2}$ for $n \ge m$. Then

$$|\mathbf{R}_n(j) - S(j)| = |S(n) - S(j)| \le |S(n) - S(m)| + |S(m) - S(j)| \le r$$

for $n \wedge j \ge m$, and so $|r_n - x| = [|\mathbf{R}_n - S|] \le r$ for $n \ge m$. Thus $r_n \longrightarrow x$. □

As a consequence of the preceding density result, we know that if $x < y$ then there is an $r \in \mathbb{Q}$ such that $x < r < y$. Therefore, if $x_n \longrightarrow x$ then for all $\epsilon \in (0, \infty) \equiv \{x \in \mathbb{R} : x > 0\}$ there is an m such that $|x_n - x| < \epsilon$ for all $n \ge m$.

The final step is to show that $\mathbb{R}$ is complete with respect to this notion of convergence, and the following lemma will allow us to do that.

Lemma A.6 *Suppose that $\{x_n : n \ge 1\} \subseteq \mathbb{R}$ and that for each $r \in \mathbb{Q}^+$ there is an m such that $|x_n - x_m| \le r$ for $n \ge m$. Then there exists a Cauchy convergent sequence $\{S_n : n \ge 1\} \subseteq \tilde{\mathbb{R}}$ such that $S_n \in x_n$ for each n.*

Proof Choose $S_n \in x_n$ for each n. Then for each $(k, n) \in (\mathbb{Z}^+)^2$ there exists $m_k \in \mathbb{Z}^+$ with the property that for any $n \ge m_k$ there is a $j_{k,n} \in \mathbb{Z}^+$ such that

$$|S_n(j) - S_{m_k}(j)| \le \frac{1}{k} \quad \text{if} \quad j \ge j_{k,n}.$$

In addition, for each $(k, n) \in (\mathbb{Z}^+)^2$ there exists an $i_{k,n} \ge j_{k,n}$ such that

$$|S_n(j) - S_n(i_{k,n})| \le \frac{1}{k} \quad \text{for all} \quad j \ge i_{k,n},$$

and, without loss in generality, we will assume that $m_{k_1} < m_{k_2}$ if $k_1 < k_2$ and $i_{k_1,n_1} \le i_{k_2,n_2}$ if $k_1 \le k_2$ and $n_1 \le n_2$.

Now define $S'_n \in \tilde{\mathbb{R}}$ so that $S'_n(k) = S_n(i_{k,n})$ for $k \in \mathbb{Z}^+$. By Lemma A.1, $S'_n \in x_n$. Furthermore, if $n \ge m_k$ and $\ell \ge k$, then

$$|S'_n(\ell) - S'_{m_k}(\ell)| \le |S_n(i_{\ell,n}) - S_{m_k}(i_{\ell,n})| + |S_{m_k}(i_{\ell,n}) - S_{m_k}(i_{k,m_k})| \le \tfrac{2}{k}.$$

Thus $\{S'_n : n \ge 1\}$ is Cauchy convergent. □

Given Lemmas A.6 and A.3, it is easy to see that if $\{x_n : n \geq 1\} \subseteq \mathbb{R}$ is Cauchy convergent in $\mathbb{R}$, then there is an $x \in \mathbb{R}$ to which it converges. Indeed, by Lemma A.6, there exist $S_n \in x_n$ such that $\{S_n : n \geq 1\}$ is Cauchy convergent in $\tilde{\mathbb{R}}$, and therefore, by Lemma A.3, there is an S to which it converges. Since this means that for each $r \in \mathbb{Q}^+$ there exists a $(m, i) \in (\mathbb{Z}^+)^2$ such that $|S_n(j) - S(j)| < \frac{r}{2}$ for $n \geq m$ and $j \geq i$, $|x_n - [S]| < r$ for $n \geq m$. With this result, we have completed the construction of a model for the real numbers.

To relate our model to more familiar ways of thinking about the real numbers, recall Exercise 1.7. It was shown there that if $D \geq 2$ is an integer and $\tilde{\Omega}$ is the set of maps $\omega : \mathbb{N} \longrightarrow \{0, \dots, D-1\}$ for which $\omega(0) \neq 0$ and $\omega(k) \neq D-1$ for infinitely many k's, then for every $x \in (0, \infty)$ there is a unique $n_x \in \mathbb{Z}$ and a unique $\omega_x \in \tilde{\Omega}$ such that $x = \sum_{k=0}^{\infty} \omega_x(k) D^{n_x - k}$. Of course, if $x < 0$ and we define $\omega_x : \mathbb{N} \longrightarrow \{0, -1, \dots, -(D-1)\}$ by $\omega_x(k) = -\omega_{|x|}(k)$, then the same equation holds. Therefore, if we define $-\omega : \mathbb{N} \longrightarrow \{0, -1, \dots, -(D-1)\}$ by $(-\omega)(k) = -\omega(k)$ for all $k \in \mathbb{N}$, then we can identify $\mathbb{R}$ with

$$\{(0, \omega^0)\} \cup \{(n, \pm\omega) : (n, \omega) \in \mathbb{Z} \times \tilde{\Omega}\},$$

where $\omega^0 : \mathbb{N} \longrightarrow \mathbb{Z}$ is given by $\omega^0(k) = 0$ for all $k \in \mathbb{N}$. In terms of Hausdorff's construction, this would correspond to choosing the partial sums $\sum_{k=0}^{j} \omega_x(k) D^{n_x - k}$ as the canonical representative from the equivalence class of x. The reason for working with equivalence classes rather than such a canonical choice of representatives is that they afford us the freedom that we needed in the proof of Lemma A.6. Namely, it is *not* true that for every choice of $S_n \in x_n$ the sequence $\{S_n : n \geq 1\}$ is Cauchy convergent in $\tilde{\mathbb{R}}$ just because $\{x_n : n \geq 1\}$ is Cauchy convergent in $\mathbb{R}$. For example, consider $\{S_n : n \geq 1\}$ where $S_n(j) = 0$ for $1 \leq j < n$ and $S_n(j) = 1$ for $j \geq n$. Then $[S_n] = 1$ for every $n \geq 1$, and yet $\{S_n : n \geq 1\}$ is not Cauchy convergent.

Index

D.W. Stroock, *A Concise Introduction to Analysis*,
DOI 10.1007/978-3-319-24469-3