Birkhäuser Advanced Texts
Basler Lehrbücher

Donald L. Cohn

Measure Theory

Second Edition

Birkhäuser

Birkhäuser

Birkhäuser Advanced Texts

For further volumes:
http://www.springer.com/series/4842

Donald L. Cohn

Measure Theory

Second Edition

Birkhäuser

Donald L. Cohn
Department of Mathematics
and Computer Science
Suffolk University
Boston, MA, USA

ISBN 978-1-4614-6955-1 ISBN 978-1-4614-6956-8 (eBook)
DOI 10.1007/978-1-4614-6956-8
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013934978

Mathematics Subject Classification (2010): 28-01, 60-01, 28C05, 28C15, 28A05, 26A39, 28A51

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.birkhauser-science.com)

To Linda, Henry, Edward, and Susan

Preface

In this new edition there are two types of changes: I have made improvements to the text of the first edition and have added some new topics.

In addition to making some corrections and reworking some arguments from the first edition, I have added an introduction before Chap. 1, in which I have said a bit about how the Lebesgue integral arose and indicated something about how the topics covered are related to one another. I hope that this will make it easier for the reader to see the structure of what he or she is studying. I have also improved the layout of the pages a bit, with the examples now easier to find.

There are a number of new topics. These main additions are the Henstock–Kurzweil integral, the Banach–Tarski paradox, and an introduction to measure-theoretic probability theory. These are, of course, supplementary to the main lines of the book, but they should give the reader a better feel for the relationship between measure theory and other parts of mathematics. As minor additions there are introductions to the Daniell integral and to the theory of liftings.

The mathematical level of the book and the background expected of the reader have not changed from the first edition.

There are several people and organizations that I would like to thank. Suffolk University's College of Arts and Sciences, together with its Department of Mathematics and Computer Science, made possible a sabbatical leave to work on this new edition. Richard Dudley and the Department of Mathematics at MIT provided office space and library access during that leave. Henry Cohn, Carl Offner, and Xinxin Jiang read and commented on parts of the manuscript. A number of people, some of whom I can no longer name, sent me useful comments on and corrections for the first edition. Ann Kostant, Tom Grasso, Kate Ghezzi, and Allen Mann, along with the production staff at Birkhäuser, were very helpful. My wife, Linda, typed parts of the manuscript, did a large amount of proofreading, and put up with my schedule as I worked on the book. I thank them all.

The Preface from the First Edition

This book is intended as a straightforward treatment of the parts of measure theory necessary for analysis and probability. The first five or six chapters form an introduction to measure and integration, while the last three chapters should provide the reader with some tools that are necessary for study and research in any of a number of directions. (For instance, one who has studied Chaps. 7 and 9 should be able to go on to interesting topics in harmonic analysis, without having to pause to learn a new theory of integration and to reconcile it with the one he or she already knows.) I hope that the last three chapters will also prove to be a useful reference.

Chapters 1 through 5 deal with abstract measure and integration theory and presuppose only the familiarity with the topology of Euclidean spaces that a student should acquire in an advanced calculus course. Lebesgue measure on $\mathbb{R}$ (and on $\mathbb{R}^d$) is constructed in Chap. 1 and is used as a basic example thereafter.

Chapter 6, on differentiation, begins with a treatment of changes of variables in $\mathbb{R}^d$ and then gives the basic results on the almost everywhere differentiation of functions on $\mathbb{R}$ (and measures on $\mathbb{R}^d$). The first section of this chapter makes use of the derivative (as a linear transformation) of a function from $\mathbb{R}^d$ to $\mathbb{R}^d$; the necessary definitions and facts are recalled, with appropriate references. The rest of the chapter has the same prerequisites as the earlier chapters.

Chapter 7 contains a rather thorough treatment of integration on locally compact Hausdorff spaces. I hope that the beginner can learn the basic facts from Sects. 7.2 and 7.3 without too much trouble. These sections, together with Sect. 7.4 and the first part of Sect. 7.6, cover almost everything the typical analyst needs to know about regular measures. The technical facts needed for dealing with very large locally compact Hausdorff spaces are included in Sects. 7.5 and 7.6.

In Chap. 8 I have tried to collect those parts of the theory of analytic sets that are of everyday use in analysis and probability. I hope it will serve both as an introduction and as a useful reference.

Chapter 9 is devoted to integration on locally compact groups. In addition to a construction and discussion of Haar measure, I have included a brief introduction to convolution on $L^1(G)$ and on the space of finite signed or complex regular Borel measures on G. The details are provided for arbitrary locally compact groups but in such a way that a reader who is interested only in second countable groups should find it easy to make the appropriate omissions.

Chapters 7 through 9 presuppose a little background in general topology. The necessary facts are reviewed, and so some facility with arguments involving topological spaces and metric spaces is actually all that is required. The reader who can work through Sects. 7.1 and 8.1 should have no trouble.

In addition to the main body of the text, there are five appendices. The first four explain the notation used and contain some elementary facts from set theory, calculus, and topology; they should remind the reader of a few things he or she may have forgotten and should thereby make the book quite self-contained. The fifth appendix contains an introduction to the Bochner integral.

Each section ends with some exercises. They are, for the most part, intended to give the reader practice with the concepts presented in the text. Some contain examples, additional results, or alternative proofs and should provide a bit of perspective. Only a few of the exercises are used later in the text itself; these few are provided with hints, as needed, that should make their solution routine.

I believe that no result in this book is new. Hence the lack of a bibliographic citation should never be taken as a claim of originality. The notes at the ends of chapters occasionally tell where a theorem or proof first appeared; most often, however, they point the reader to alternative presentations or to sources of further information.

The system used for cross-references within the book should be almost self-explanatory. For example, Proposition 1.3.5 and Exercise 1.3.7 are to be found in Sect. 1.3 of Chap. 1, while C.1 and Theorem C.8 are to be found in Appendix C.

There are a number of people to whom I am indebted and whom I would like to thank. First there are those from whom I learned integration theory, whether through courses, books, papers, or conversations; I won't try to name them, but I thank them all. I would like to thank R.M. Dudley and W.J. Buckingham, who read the original manuscript, and J.P. Hajj, who helped me with the proofreading. These three read the book with much care and thought and provided many useful suggestions. (I must, of course, accept responsibility for ignoring a few of their suggestions and for whatever mistakes remain.) Finally, I thank my wife, Linda, for typing and providing editorial advice on the manuscript, for helping with the proofreading, and especially for her encouragement and patience during the years it took to write this book.

Boston, MA, USA Donald L. Cohn

Contents

Introduction

In this introduction we

- briefly review the Riemann integral as studied in calculus and elementary analysis,
- sketch how some difficulties with the Riemann integral led to the Lebesgue integral, and
- outline the main topics in this book and note how they relate to the Riemann and Lebesgue integrals.

The Riemann Integral—Darboux's Definition

Let $[a,b]$ be a closed bounded interval. A *partition* of $[a,b]$ is a finite sequence $\{a_i\}_{i=0}^k$ of real numbers such that

$$a = a_0 < a_1 < \cdots < a_k = b.$$

Sometimes we will call the values a_i the *division points* of the partition. We will generally denote a partition by a symbol such as $\mathscr{P}$.

Suppose that f is a bounded real-valued function on $[a,b]$ and that $\mathscr{P}$ is a partition of $[a,b]$, say with division points $\{a_i\}_{i=0}^k$. For $i = 1, \ldots, k$ define numbers m_i and M_i by $m_i = \inf\{f(x) : x \in [a_{i-1}, a_i]\}$ and $M_i = \sup\{f(x) : x \in [a_{i-1}, a_i]\}$. Then the *lower sum* $l(f,\mathscr{P})$ corresponding to f and $\mathscr{P}$ is defined to be $\sum_{i=1}^k m_i(a_i - a_{i-1})$. Similarly, the *upper sum* $u(f,\mathscr{P})$ corresponding to f and $\mathscr{P}$ is defined to be $\sum_{i=1}^k M_i(a_i - a_{i-1})$. See Fig. 1 below.

Since f is bounded, there are real numbers m and M such that $m \le f(x) \le M$ holds for each x in $[a,b]$. Then each lower sum of f satisfies

$$l(f,\mathscr{P}) = \sum_{i=1}^k m_i(a_i - a_{i-1}) \le \sum_{i=1}^k M(a_i - a_{i-1}) = M(b-a),$$

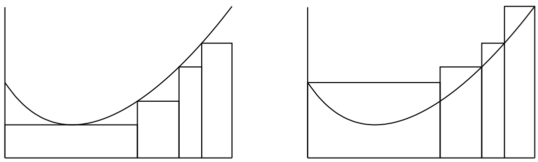

Fig. 1 A lower sum and an upper sum

and so the set of lower sums of f is bounded above, in fact by $M(b-a)$. It follows that the set of lower sums has a supremum (a least upper bound); this supremum is called the lower integral of f over $[a,b]$ and is denoted by $\underline{\int}_a^b f$. A similar argument shows that the set of upper sums of f is bounded below, and so one can define the upper integral of f, written $\overline{\int}_a^b f$, to be the infimum (the greatest lower bound) of the set of upper sums. It is not difficult to show (see Sect. 2.5 for details) that $\underline{\int}_a^b f \le \overline{\int}_a^b f$. If $\underline{\int}_a^b f = \overline{\int}_a^b f$, then f is said to be *Riemann integrable* on $[a,b]$, and the common value of $\underline{\int}_a^b f$ and $\overline{\int}_a^b f$ is called the *Riemann integral* of f over $[a,b]$ and is denoted by $\int_a^b f$ or $\int_a^b f(x)\,dx$.

The Riemann Integral—Riemann's Definition

It is sometimes useful to view Riemann integrals as limits of what are called Riemann sums. For this we need a couple of definitions. A *tagged partition* of an interval $[a,b]$ is a partition $\{a_i\}_{i=0}^k$ of $[a,b]$, together with a sequence $\{x_i\}_{i=1}^k$ of numbers (called *tags*) such that $a_{i-1} \le x_i \le a_i$ holds for $i = 1, \dots, k$. (In other words, each tag x_i must belong to the corresponding interval $[a_{i-1}, a_i]$.) As with partitions, we will often denote a tagged partition by a symbol such as $\mathscr{P}$.

The *mesh* $\|\mathscr{P}\|$ of a partition (or of a tagged partition) $\mathscr{P}$ is defined by $\|\mathscr{P}\| = \max_i(a_i - a_{i-1})$, where $\{a_i\}$ is the sequence of division points for $\mathscr{P}$. In other words, the mesh of a partition is the length of the longest of its subintervals.

The *Riemann sum* $\mathscr{R}(f,\mathscr{P})$ corresponding to the function f and the tagged partition $\mathscr{P}$ is defined by

$$\mathscr{R}(f,\mathscr{P}) = \sum_{i=1}^k f(x_i)(a_i - a_{i-1}).$$

Then, according to Riemann's definition, the function f is integrable over $[a,b]$ if there is a number L (which will be the value of the integral) such that

$$\lim_{\mathscr{P}} \mathscr{R}(f,\mathscr{P}) = L,$$

where the limit is taken as the mesh of $\mathscr{P}$ approaches 0. If we express this in terms of ε's and δ's, we see that the function f is Riemann integrable, with integral L, if for every positive ε there is a positive δ such that $|\mathscr{R}(f,\mathscr{P}) - L| < \varepsilon$ holds for each tagged partition $\mathscr{P}$ of $[a,b]$ that satisfies $\|\mathscr{P}\| < \delta$.

Darboux's and Riemann's definitions are equivalent:[1] they give exactly the same classes of integrable functions, with the same values for the integrals (see Proposition 2.5.7).

Another standard result is that every continuous function on $[a,b]$ is Riemann integrable; see Example 2.5.2 (or, for a somewhat stronger result, see Theorem 2.5.4).

The final thing to recall is the fundamental theorem of calculus (see Exercise 2.5.6 for a sketch of its proof):

Theorem 1 (The Fundamental Theorem of Calculus). *Suppose that $f\colon [a,b] \to \mathbb{R}$ is continuous and that $F\colon [a,b] \to \mathbb{R}$ is defined by $F(x) = \int_a^x f(t)\,dt$. Then F is differentiable at each x in $[a,b]$ and its derivative is given by $F'(x) = f(x)$.*

From Riemann to Lebesgue

In many situations involving integrals (for example, when integrating an infinite series term by term or when differentiating under the integral sign), it is necessary to be able to reverse the order of taking limits and evaluating integrals—that is, to be able to say things like

$$\int_a^b \lim_n f_n(x)\,dx = \lim_n \int_a^b f_n(x)\,dx.$$

Thus one needs to have theorems of the following sort:

Theorem 2. *Suppose that $\{f_n\}$ is a sequence of integrable functions on the interval $[a,b]$ and that f is a function such that $\{f_n\}$ converges to f in a suitable[2] way. Then f is integrable and*

$$\int_a^b f(x)\,dx = \lim_n \int_a^b f_n(x)\,dx.$$

In elementary analysis courses one sees that Theorem 2 is valid for the Riemann integral if by "converges to f in a suitable way," we mean "converges uniformly to f" (see Exercise 2.5.7). On the other hand, if we do not assume uniform

---

[1]The reader may well be asking why people consider two definitions of the Riemann integral. The general answer is that Darboux's definition is simpler and more elegant, while Riemann's is useful for various calculations of limits (see, for example, Exercise 2.5.8). For our purposes, the Darboux approach makes our discussion of the relationship between the Riemann and Lebesgue integrals simpler, while the Riemann approach is more closely related to the Henstock–Kurzweil and McShane integrals (see Appendix H).

[2]The problem is, of course, to figure out what "suitable" might mean and to define the integral in such a way that theorems like this one will be applicable in many situations.

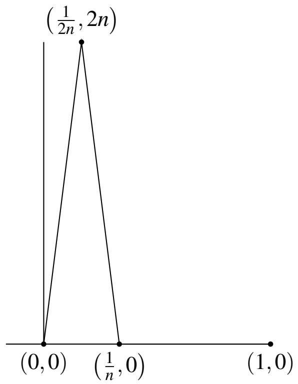

Fig. 2 Function defined in Example 3

convergence of $\{f_n\}$ to f, but only pointwise[3] convergence, then, as we see in the following examples, Theorem 2 may fail.

Example 3. For each positive integer n let f_n be the piecewise linear function on $[0,1]$ whose graph is made up of three line segments, connecting the points $(0,0)$, $(\frac{1}{2n}, 2n)$, $(\frac{1}{n}, 0)$, and $(1,0)$. See Fig. 2. Then for each n the triangle formed by the graph of f_n and the x-axis has area 1, and so f_n satisfies $\int_0^1 f_n(x)\,dx = 1$. Furthermore, for each x in $[0,1]$ we have $\lim_n f_n(x) = 0$. Thus $\lim_n \int_0^1 f_n(x)\,dx = 1$ but $\int_0^1 \lim_n f_n(x)\,dx = 0$, and the conclusion of Theorem 2 fails for the sequence $\{f_n\}$. □

The failure of the conclusion of Theorem 2 in the preceding example comes from the fact that the sequence $\{f_n\}$ is not uniformly bounded—that is, from the fact that there is no constant M such that $|f_n(x)| \le M$ holds for all n and x. Next let us look at an example in which the functions f_n are uniformly bounded, in fact, in which we have $0 \le f_n(x) \le 1$ for all n and all x, and yet the conclusion to Theorem 2 fails.

Example 4. Recall that the set of rational numbers is countable (see A.6). Hence we can choose an enumeration $\{x_n\}$ of the rational numbers in the interval $[0,1]$ (that is, a sequence whose members are the rational numbers in $[0,1]$, with each rational in that interval occurring exactly once in the sequence). For each n define a function $f_n \colon [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} 1 & \text{if } x \in \{x_1, x_2, \dots, x_n\}, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

[3]Recall that $\{f_n\}$ converges pointwise to f on $[a,b]$ if $\lim_n f_n(x) = f(x)$ for each x in $[a,b]$.

Thus $f_n(x)$ has value 1 for n values of x (namely for $x_1, \ldots, x_n$) and has value 0 otherwise. It is easy to check that for each n, all the lower sums of f_n are 0 and hence that the lower integral $\underline{\int}_0^1 f_n$ is 0. On the other hand, it is not hard to construct, for each n and each positive δ, a partition $\mathscr{P}$ of $[0,1]$ in which each of $x_1, x_2, \ldots, x_n$ is in the interior of some subinterval that belongs to $\mathscr{P}$ and has length at most δ/n. It follows that $u(f_n, \mathscr{P}) \leq \delta$. Since this can be done for each positive δ, it follows that the upper integral $\overline{\int}_0^1 f_n$ is also 0. Consequently f_n is Riemann integrable over $[0,1]$ and $\int_0^1 f_n(x)\,dx = 0$.

For each x let us consider the behavior of the sequence $\{f_n(x)\}$. If x is rational, then $f_n(x) = 1$ for all large n, while if x is irrational, then $f_n(x) = 0$ for all n. Thus $\{f_n(x)\}$ converges pointwise to the function $f\colon [0,1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational and belongs to } [0,1], \text{ and} \\ 0 & \text{if } x \text{ is irrational and belongs to } [0,1]. \end{cases}$$

Since the rationals are dense in $[0,1]$, as are the irrationals, it follows that every lower sum for f has value 0 and every upper sum for f has value 1. Thus the lower and upper integrals of f are given by $\underline{\int}_0^1 f = 0$ and $\overline{\int}_0^1 f = 1$, and f is not Riemann integrable. Thus the conclusion of Theorem 2 fails for this example. □

Example 5. It may seem that the difficulty in the previous example comes from the fact that the functions f_n fail to be continuous. However, one can also produce a sequence $\{f_n\}$ such that

(a) each f_n is continuous,
(b) $0 \leq f_n(x) \leq 1$ holds for each n and each x, and
(c) $\{f_n\}$ converges pointwise to a function that is not Riemann integrable.

(See Exercise 2.5.4.) □

The questions involved in making Theorem 2 precise were important unresolved issues in the late nineteenth century; they arose, for example, in the study of Fourier series.

In the early twentieth century, Lebesgue defined a new integral, which he used to give very useful answers to questions of the sort discussed above. For example, Lebesgue showed that Theorem 2, when formulated in terms of his new integral, holds for pointwise convergence of the sequence $\{f_n\}$, subject only to some rather natural boundedness conditions on that sequence (see the dominated convergence theorem, Theorem 2.4.5). It is hard to overemphasize the simplicity and ease of application of the limit theorems for the Lebesgue integral.

Let us briefly sketch how the Lebesgue integral is defined. For simplicity, we will for now restrict our attention to functions $f\colon [a,b] \to \mathbb{R}$ that are nonnegative and bounded (those assumptions are in no way necessary). So let c be a positive number such that $0 \leq f(x) < c$ holds for each x in $[a,b]$. As we have seen, the definition of the Riemann integral deals with partitions of the interval $[a,b]$, that is, of the domain of f. One way of defining the Lebesgue integral deals with partitions of the range of f, rather than of the domain. So suppose that $\mathscr{P}$ is a partition of $[0,c]$, say given by a sequence of $\{a_i\}_{i=0}^k$ of dividing points. For $i = 1, \ldots, k$ define A_i by

$$A_i = \{x \in [a,b] \colon f(x) \in [a_{i-1}, a_i)\}. \tag{1}$$

(Note that the sets A_i are not necessarily subintervals of $[a,b]$—they can also be empty, unions of finite collections of subintervals, or even more complicated sets.) Let us consider the sum $s(f,\mathscr{P})$ given by

$$s(f,\mathscr{P}) = \sum_{i=1}^{k} a_{i-1}\operatorname{meas}(A_i), \tag{2}$$

where $\operatorname{meas}(A_i)$ is the size, in a sense still to be defined, of the set A_i. Subject to the condition that the function f must be simple enough that $\operatorname{meas}(A_i)$ makes sense for all sets A_i as defined by (1), the Lebesgue integral of f is defined to be the supremum of the set of all sums of the form (2), where these sums are considered for all partitions $\mathscr{P}$ of the interval $[0,c]$. (One can check that this does not depend on the value of c, as long as it is large enough that $f(x) < c$ holds for all x.)

Now let us survey some of the contents of this book.

The first issue that needs resolving is the meaning of the expression $\operatorname{meas}(A_i)$ that occurs in Eq. (2). That is the goal of Chap. 1, which begins with the question of how to describe and organize the subsets of $\mathbb{R}$ whose size can reasonably be measured (that is, the *measurable* sets) and then continues with the question of how to measure the sizes of those subsets (the study of Lebesgue measure and of more general measures). Since it is useful to consider integration not just for functions defined on $\mathbb{R}$ or on subintervals of $\mathbb{R}$ but also in more general settings, including $\mathbb{R}^d$, some of the discussion in Chap. 1 is rather abstract. This abstractness does not add much to the level of difficulty of the chapter.

Appendix G is in some sense a continuation of Chap. 1. It gives an exposition of the Banach–Tarski paradox, which is a very famous result that quite vividly shows that Lebesgue measure on $\mathbb{R}^3$ cannot be extended in any reasonable way to all the subsets of $\mathbb{R}^3$. (Appendix G is deeper than Chap. 1 and requires more background on the reader's part.)

The main objective of Chap. 2 is the definition of the Lebesgue integral. Section 2.1 deals with measurable functions, those functions that are tame enough that the sets A_i in Eq. (2) are measurable. Section 2.2 introduces properties that hold *almost everywhere* and in particular considers convergence almost everywhere, which can often be used in place of pointwise convergence. The integral is finally defined in Sect. 2.3, and the basic limit theorems for the integral are proved in Sect. 2.4.

Chapter 3 deals more deeply with limits and convergence in integration theory, while Chap. 4 deals with measures that have signed or complex values and with relationships between measures.

In multivariable calculus courses one learns how to calculate integrals over subsets of $\mathbb{R}^d$ by repeatedly calculating one-dimensional integrals. Chapter 5 deals with such matters for the Lebesgue integral. Section 6.1 deals with another aspect of integration on $\mathbb{R}^d$, namely with change of variable in integrals over subsets of $\mathbb{R}^d$.

The fundamental theorem of calculus (Theorem 1 above) relates Riemann integrals to derivatives. Such relationships for the Lebesgue integral are discussed in the last two sections of Chap. 6.

In the discussion above of Chap. 1 we noted that our treatment of measures and measurable sets is fairly general. This generality is useful for a number of applications, such as to cases where integration on locally compact topological spaces is needed (see Chaps. 7 and 9) and to the study of probability theory (see Chap. 10 for a brief introduction to the application of measure theory to probability theory).

Many deeper questions about measurable sets and functions arise naturally. Some useful and classical results along these lines are given in Chap. 8.

Let us return for a moment to the second of our definitions of the Riemann integral, the one expressed in terms of limits of Riemann sums. In the second half of the twentieth century Henstock and Kurzweil gave what may seem to be a small modification of this definition. The resulting integral is known as the Henstock–Kurzweil integral or the generalized Riemann integral. Although their definition seems very simple, their integral (for functions on $\mathbb{R}$) turns out to be more general than the Lebesgue integral and to have what is in some ways a more natural relationship to derivatives. See Appendix H for an introduction to the Henstock–Kurzweil integral.