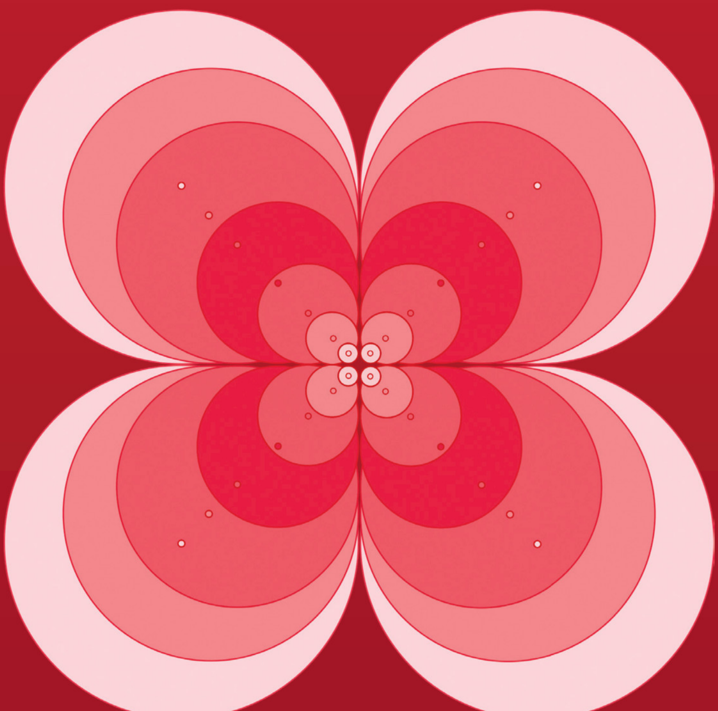
Counterexamples in Measure and Integration

René L. Schilling and
Franziska Kühn



COUNTEREXAMPLES IN MEASURE AND INTEGRATION

Often it is more instructive to know 'what can go wrong' and to understand 'why a result fails' than to plod through yet another piece of theory. In this text, the authors gather more than 300 counterexamples – some of them both surprising and amusing – showing the limitations, hidden traps and pitfalls of measure and integration. Many examples are put into context, explaining the relevant parts of the theory, and pointing out further reading.

The text starts with a self-contained, non-technical overview on the fundamentals of measure and integration. A companion to the successful undergraduate textbook *Measures, Integrals and Martingales*, it is accessible to advanced undergraduate students, requiring only modest prerequisites. More specialized concepts are briefly summarized at the beginning of each chapter, allowing for self-study as well as supplementary reading for any course covering measures and integrals. For researchers, the text provides ample examples and warnings as to the limitations of general measure theory.

RENÉ L. SCHILLING is Professor of Probability Theory at Technische Universität Dresden. His research focuses on stochastic analysis and the theory of stochastic processes.

FRANZISKA KÜHN is Research Assistant at Technische Universität Dresden, where she finished her Ph.D. in 2016. She is interested in the interplay of probability theory and analysis, with a focus on jump processes and non-local operators.

COUNTEREXAMPLES IN MEASURE AND INTEGRATION

RENÉ L. SCHILLING
Technische Universität Dresden

FRANZISKA KÜHN
Technische Universität Dresden

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781316519134
DOI: 10.1017/9781009003797

First published 2021

Printed in the United Kingdom by TJ Books Limited, Padstow Cornwall

A catalogue record for this publication is available from the British Library.

ISBN 978-1-316-51913-4 Hardback
ISBN 978-1-009-00162-5 Paperback

Contents

Preface

A **counterexample** /ˈkaʊntərɪɡˌzɑːmpl/ is *an example that opposes or contradicts an idea or theory.*[1] It is fair to say that the word 'counterexample' is not too common in everyday language, but rather a concept from philosophy and, of course, mathematics. In mathematics, there are proofs and examples, and while an example, say, of some $x \in A$ satisfying $x \in B$ does not prove $A \subseteq B$, the *counter*example of some $x_0 \in B$ such that $x_0 \notin A$ disproves $A \subseteq B$; in other words, it proves that $A \subseteq B$ does not hold. This observation shows that there is no sharp distinction between example and counterexample, and we do not give a definition of what a counterexample should or could be (you may want to consult Lakatos [94] instead), but assume the more pragmatic point of view of a working mathematician. If we want to solve a problem, we look at the same time for a proof and for counterexamples which help us to capture and delineate the subject matter.

The same is also true for the student of mathematics, who will gain a better understanding of a theorem or theory if he knows its limitations – which may be expressed in the form of counterexamples. The present collection of (counter-)examples grew out of our own experience, in the classroom and on `stackexchange.com`, where we are often asked after the 'how' and 'why' of many a result. This explains the wide range of examples, from the fairly obvious to rather intricate constructions. The choice of the examples reflects, naturally, our own taste. We decided to include only those counterexamples which could be dealt with in a couple of pages (or less) and which are not too pathological – one can, indeed, destroy almost anything by the choice of the underlying topology. We intend the present volume as a companion to our textbook *Measures, Integrals and Martingales* [MIMS], which means that most examples are from elementary measure and integration, not touching on integration on

[1] Oxford dictionaries `https://en.oxforddictionaries.com/definition/counterexample`, accessed 11-May-2019.

groups (Haar measure) or on really deep axiomatic issues (e.g. as in descriptive set theory, see Kechris [89], and the advanced constructive theory of functions, see Kharazishvili [91, 92]).

This book is intended as supplementary reading for a course in measure and integration theory, or for seminars and reading courses where students can explore certain aspects of the theory by themselves. Where appropriate, we have added comments putting the example into context and pointing the reader to further literature. We think that this book will also be useful for lecturers and tutors in teaching measure and integration, and for researchers who may discover new and sometimes unexpected phenomena. Readers are assumed to have basic knowledge of functional analysis, point-set topology and, of course, measure and integration theory. For novices, there is a panorama of measure and integration which gives a non-technical overview on the subject and can serve, to some extent, as a first introduction. The overall presentation is as self-contained as possible; in order to make the text easy to access, we use only a few standard references – Schilling [MIMS] and Bogachev [19] for measure and integration, Rudin [151] and Yosida [202] for functional analysis, and Willard [200] and Engelking [53] for topology.

Some of the counterexamples are famous, many are more or less well known, and a few are of our own making. When we could trace the origin of an example, we have given references and attached names, but most entries are 'standard' examples which seem to have been in the public domain for ages; having said this, we acknowledge a huge debt to many anonymous authors and we do apologize if we have failed to give proper credit. The three classic counterexample books by Gelbaum & Olmsted [65], Steen & Seebach [172], and Stoyanov [180] were both inspiration and encouragement. We hope that this book lives up to their high standards.

It is a pleasure to acknowledge the interest and skill of our editor, Roger Astley, in the preparation of this book and Cambridge University Press for the excellent book design. Many colleagues have contributed to this text with comments and suggestions, in particular M. Auer, R. Baumgarth, G. Berschneider, N.H. Bingham – for the famous full red-ink treatment, C.-S. Deng, D.E. Edmunds and C. Goldie – for most helpful discussions, Y. Ishikawa, N. Jacob – for access to his legendary library, Y. Mishura and N. Sandrić. We thank our colleagues and friends who suffered for quite a while from our destructive search for counterexamples (*Do you know an example of a measure which fails to … ?*), strange functions and many outer-worldly excursions – and our families who have us back in real life.

User's Guide

This book is not intended for linear reading – although this might well be possible – but invites the reader to browse, to read selectively and to look things up. We have, therefore, organized the material in self-contained chapters which treat different aspects of measure and integration theory. We assume that the reader has a basic knowledge of abstract measure and integration; the outline given in the 'panorama' (Chapter 1) is intended to refresh the reader's memory, to fix notation and to give a first non-technical introduction to the subject. The cross-reference [☞ *n.m*] appearing in the margin points towards essential counterexamples to the (positive) result at hand. Some supplementary material which is not always part of the mathematical curriculum is collected in Chapter 2; look it up once you need it. ☞ *n.m*

Cross-referencing. Throughout the text, [☞ *n.m*] and Example *n.m* refers to counterexample *m* in Chapter *n*. Theorem *n.m*, Definition *n.m*, etc. point to the respective theorem, definition, etc. in the 'panorama' (Chapter 1) or the 'refresher' (Chapter 2). Equation *m* in Chapter *n* is denoted by (*n.m*). At the beginning of each chapter, we recall more specialized results and definitions which are particular to that chapter; these are numbered locally as 5*A*, 5*B*, 5*C*, … (for Chapter 5, say) and they are mostly used within that chapter. Theorems, lemmas and corollaries may also appear in a counterexample; if needed, we use again local numbering 1, 2, 3, … .

Finding stuff. Following Gelbaum & Olmsted [65] we have organized the examples by theme and all counterexamples are listed in the list of contents by (hopefully) meaningful names. We begin with examples on Riemann integration (Chapter 3) and move on to various aspects of the (abstract) Lebesgue integral (Chapters 4–19). The chapters on Lebesgue integration follow 'The way of integration' (alluding to Fig. 1.3 in Chapter 1), i.e. beginning with measurable sets and σ-algebras to set functions, measurable functions, to integrals and theorems on integration. The subject index helps to find definitions, theorems and concepts, but it does not refer to specific counterexamples.

Notation. We tried to avoid specialized notation and we use commonly accepted standard notation, e.g. as in [MIMS]. The following list is intended to aid cross-referencing, so notation that is specific to a single section is generally not listed; numbers following entries are page numbers.

Unless otherwise stated, binary operations between functions such as $f \pm g$, $f \cdot g$, $f \wedge g$, $f \vee g$, comparisons $f \leqslant g$, $f < g$ or limiting relations $f_n \xrightarrow[n\to\infty]{} f$, $\lim_n f_n$, $\liminf_n f_n$, $\limsup f_n$, $\sup_i f_i$ or $\inf_i f_i$ are always understood pointwise.

General notation

positive	always in the sense $\geqslant 0$
negative	always in the sense $\leqslant 0$
increasing	$x \leqslant y \implies f(x) \leqslant f(y)$
decreasing	$x \leqslant y \implies f(x) \geqslant f(y)$
countable	finite or countably infinite
$\mathbb{N}$	natural numbers: $1, 2, 3, \ldots$
$\mathbb{N}_0$	positive integers: $0, 1, 2, \ldots$
$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$	integer, rational, real, complex numbers
$\overline{\mathbb{R}}$	$[-\infty, +\infty]$ (two-point compactification), 11, 38
$\inf\emptyset$, $\sup\emptyset$	$\inf\emptyset = +\infty$, $\sup\emptyset = -\infty$
$a \vee b$, $a \wedge b$	$\max\{a, b\}$, $\min\{a, b\}$
$\gcd(\cdot, \cdot)$	greatest common divisor
$\aleph_0$	cardinality of $\mathbb{N}$, 44
$\mathfrak{c}$	cardinality of $\mathbb{R}$, 44
ω_0	first infinite ordinal, ordinal number of $\mathbb{N}$, 44
ω_1	first uncountable ordinal, 45, 46
$\Omega = [0, \omega_1]$	ordinal space, 45, 46
$\Omega_0 = [0, \omega_1)$	countable ordinals, 45, 46

Sets and set operations

$A \mathbin{\dot\cup} B$	union of disjoint sets
$A \vartriangle B$	$(A \setminus B) \mathbin{\dot\cup} (B \setminus A)$
A^c	complement of A
$\overline{A}$	closure of A, 37
A°	open interior of A, 37
$A_n \uparrow A$	$A_n \subseteq A_{n+1}$, $A = \bigcup_n A_n$
$A_n \downarrow A$	$A_n \supseteq A_{n+1}$, $A = \bigcap_n A_n$
$\#A$	cardinality of A
$B_r(x)$	open (metric) ball $\{y;\ d(x,y) < r\}$

Families of sets

$\mathscr{A}, \mathscr{B}, \mathscr{C}$	generic families of sets
$\mathscr{A}^*$	μ^* measurable sets, 28 completion, 9
$\mathscr{A} \otimes \mathscr{B}$	product σ-algebra, 10, 21
$\mathscr{B}(X)$	Borel sets in X, 9
$\mathscr{L}(X)$	Lebesgue sets in X, 10
$\mathscr{O}(X)$	open sets in X, 36
$\mathscr{P}(X)$	all subsets of X
$\sigma(\mathscr{F})$	σ-algebra generated by $\mathscr{F}$, 9
$\sigma(\phi)$, $\sigma(\phi_i, i \in I)$	σ-algebra generated by the map(s) ϕ, resp. ϕ_i, 9

Measures and integrals

μ, ν	generic (positive) measures
μ_*, μ^*	inner and outer measure, 182, 100
δ_x	Dirac measure in x, 8
λ, λ^d	Lebesgue measure, 9
$\zeta, \zeta_X, \#(\cdot)$	counting measure on X, 8
$\mu \circ f^{-1}, f_*\mu$	image or push-forward measure, 8, 24

$\mu \times \nu$	product of measures, 10, 21
$\mu * \nu$	convolution, 26
$\mu \ll \nu$	absolute continuity, 29
$\mu \perp \nu$	singular measures, 29
$\frac{d\nu}{d\mu}$	Radon–Nikodým derivative, 29
$\operatorname{supp} \mu$	support of a measure, 123
$\overline{\int}$, $\underline{\int}$	upper, lower R-integral, 2

Functions and spaces

$\mathbb{1}_A$	$\mathbb{1}_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$
$\operatorname{sgn}(x)$	$\mathbb{1}_{(0,\infty)}(x) - \mathbb{1}_{(-\infty,0)}(x)$
$f(A)$	$\{f(x)\,;\ x \in A\}$
$f^{-1}(\mathscr{B})$	$\{f^{-1}(B)\,;\ B \in \mathscr{B}\}$
f^+	$\max\{f(x), 0\}$ positive part
f^-	$-\min\{f(x), 0\}$ negative part
$\{f \in B\}$	$\{x\,;\ f(x) \in B\}$
$\{f \geqslant \lambda\}$	$\{x\,;\ f(x) \geqslant \lambda\}$, etc.
$f * g$	convolution, 26
$\operatorname{supp} f$	support $\overline{\{f \neq 0\}}$
$C(X)$	continuous functions on X
$C_b(X)$	bounded — —
$C_c(X)$	— — with compact support
L^p, L^∞	Lebesgue spaces, 15
$\\|f\\|_p, \\|f\\|_{L^p}$	$(\int \|f\|^p \, d\mu)^{1/p}$, $p < \infty$
$\\|f\\|_\infty, \\|f\\|_{L^\infty}$	$\operatorname{esssup} f :=$ $\inf\{c\,;\ \mu\{\|f\| \geqslant c\} = 0\}$, 14

List of Topics and Phenomena

Topic	Possible consequence	Example
μ is not finite	▶ μ not continuous from above ▶ range of μ not closed ▶ Jensen's inequality does not hold ▶ $p \leqslant q \not\Rightarrow L^q \subseteq L^p$ ▶ no series test for integrability ▶ Egorov's theorem fails ▶ convergence in probability $\not\Rightarrow$ convergence in measure	[☞ 5.10] [☞ 6.17] [☞ 18.1] [☞ 10.5] [☞ 11.11] [☞ 11.5]
μ is not σ-finite	▶ no unique product measure ▶ Fubini's and Tonelli's theorem fail ▶ Radon–Nikodým's theorem fails ▶ Lebesgue's decomposition theorem fails ▶ there is no positive integrable function ▶ limits in probability not unique ▶ $f_n \to f$ in probability $\not\Rightarrow f_{n_k} \to f$ a.e. ▶ L^p, $1 \leqslant p < \infty$, not separable ▶ $(X, \mathscr{A}^*, \mu^*\vert_{\mathscr{A}^*}) \neq$ completion of $(X, \mathscr{A}, \mu)$ ▶ trace of a regular measure not regular	[☞ 16.1] [☞ 16.8–16.18] [☞ 17.1–17.3] [☞ 17.11] [☞ 10.20] [☞ 11.6] [☞ 11.7] [☞ 18.14] [☞ 9.9] [☞ 9.22]
μ does not have the finite subset property	▶ $\int_A f\,d\mu = \int_A g\,d\mu$ for all $A \not\Rightarrow f = g$ a.e. ▶ $\exists f \in L^\infty$ s.t. $\Lambda_f(g) = \int fg\,d\mu$, $g \in L^1$, satisfies $\Vert\Lambda_f\Vert < \Vert f\Vert_{L^\infty}$ ▶ $\int \vert fg\vert\,d\mu \leqslant C\Vert g\Vert_{L^q} \not\Rightarrow f \in L^p$	[☞ 10.23] [☞ 18.21] [☞ 18.19]
μ is not locally finite	▶ $C_b(X) \cap L^p(\mu)$ not dense in $L^p(\mu)$, $1 \leqslant p < \infty$	[☞ 18.16]
μ is not regular	▶ Lusin's theorem fails ▶ $C_b(X) \cap L^p(\mu)$ not dense in $L^p(\mu)$, $1 \leqslant p < \infty$ ▶ there exists $\nu \neq \mu$ s.t. $\int f\,d\mu = \int f\,d\nu$ for all $f \in C_c(X)$	[☞ 13.16] [☞ 18.16] [☞ 18.26, 18.27]

Topic	Possible consequence	Example
X is not separable	▶ $\mathcal{B}(X)$ not generated by open balls ▶ $\mathcal{B}(X)$ not countably generated ▶ $\operatorname{supp}\mu \neq$ smallest closed set F such that $\mu(X \setminus F) = 0$ ▶ $\mu(X) \neq \mu(\operatorname{supp}\mu)$ ▶ finite measures not tight	[☞4.14] [☞4.10] [☞6.2] [☞6.3] [☞5.26]
X is not a metric space	▶ compact sets not Borel ▶ fewer Baire sets than Borel sets ▶ pointwise limits of measurable functions not measurable ▶ finite measures not outer regular ▶ inner compact regular $\not\Rightarrow$ inner regular	[☞4.15] [☞4.24] [☞8.16] [☞9.18] [☞9.18]
X is not σ-compact	▶ locally finite $\not\Rightarrow$ σ-finite ▶ inner regular $\not\Rightarrow$ inner compact regular	[☞5.16] [☞9.20]
X is not locally convex	▶ only trivial dual space ▶ no Bochner integral	[☞18.11] [☞18.34]
X has cardinality $> \mathfrak{c}$	▶ the diagonal is not in $\mathcal{P}(X) \otimes \mathcal{P}(X)$ ▶ $\mathcal{B}(X) \otimes \mathcal{B}(X) \neq \mathcal{B}(X \times X)$ ▶ metric not jointly measurable with respect to $\mathcal{B}(X) \otimes \mathcal{B}(X)$	[☞15.9] [☞15.6] [☞15.10]
$\mathcal{A}$ is too small, e.g. discrete	▶ 'few' measurable functions $f : X \to \mathbb{R}$ ▶ factorization lemma fails	[☞8.2] [☞8.20]
$\mathcal{A}$ is too big, e.g. discrete	▶ 'many' measurable functions $f : X \to \mathbb{R}$ ▶ 'few' non-atomic measures	[☞8.1] [☞6.15]
$\mathcal{A}$ not countably generated	▶ two-valued measures which are not a point mass	[☞6.10]
role of 'small' sets	▶ Lebesgue null sets may be uncountable/of second category ▶ $B + B = \mathbb{R}$ for a Lebesgue null set B ▶ $2^{\mathfrak{c}}$ many Lebesgue sets but 'only' $\mathfrak{c}$ many Borel sets ▶ $f' = 0$ a.e. $\not\Rightarrow$ f constant ▶ f a.e. continuous $\not\Rightarrow$ $f = g$ a.e. for g continuous ▶ $\mu_n \to \mu$ weakly $\not\Rightarrow$ $\mu_n(B) \to \mu(B)$ for all B ▶ support of a probability measure may have measure 0	[☞7.4, 7.9] [☞7.27] [☞4.20] [☞2.6, 14.5] [☞13.1] [☞19.5] [☞6.3]
lack of countability	▶ many theorems fail for nets, e.g. classical convergence theorems, Egorov's and Lévy's continuity theorem ▶ uncountable supremum of measurable functions are not measurable ▶ $\mathcal{B}(X)^{\otimes I}$ is 'small' for I uncountable	[☞12.9] [☞11.12, 19.12] [☞8.18] [☞15.8, 4.17]

Topic	Possible consequence	Example
	▶ projective limit of consistent family may not exist	[☞ 16.21]
	▶ $t \longmapsto f(t,x)$ cts. $\forall x$ $\not\Rightarrow t \longmapsto \int f(t,x)\,\mu(dx)$ cts.	[☞ 14.12]
lack of uniform integrability	▶ $(f_n)_{n\in\mathbb{N}} \subseteq L^1$, $f_n \to 0$ a.e. $\not\Rightarrow \int f_n \to 0$	[☞ 12.1, 12.7]
	▶ $f_n \to f$, $f_n' \to g$ everywhere $\not\Rightarrow f' = g$ a.e.	[☞ 14.8]
	▶ $f_n \to f$ in probability $\not\Rightarrow f_{n_k} \to f$ in measure	[☞ 11.8]
	▶ sequential weak compactness fails	[☞ 18.29]
L^1, L^∞ are special	▶L^∞ separable if, and only if, $\dim L^\infty < \infty$	[☞ 18.15]
	▶$(L^1)^* \supsetneqq L^\infty$	[☞ 18.21, 18.22]
	▶$(L^\infty)^* \supsetneqq L^1$	[☞ 18.24]
	▶ not uniformly convex	[☞ 18.32]
⚠ atom	▶ comparison: different definitions of atom	[☞ 6.11]
⚠ absolute continuity	▶ comparison: different definitions of absolute continuity	[☞ 17.4]
⚠ convergence in measure	▶ comparison: convergence in measure vs. in probability	[☞ p. 20 Fig. 1.4 pp. 221, 226–227]
⚠ weak convergence	▶ weak convergence of measures is not weak convergence in the sense of functional analysis	[☞ p. 371, 19.7]
⚠ Baire σ-algebra	▶ comparison of different definitions of Baire sets	[☞ 4.23]

1

A Panorama of Lebesgue Integration

The idea of measuring area and volume by infinitesimal (exhaustion) methods was already known to the ancient Greeks. This may be seen as the first example of 'integration'. The precursor of our modern notion of integration begins with the creation of the infinitesimal calculus by Newton and Leibniz. For Newton, the derivative was the primary operation of calculus and the integral was just the primitive, i.e. the antiderivative. Leibniz followed a more geometric approach, defining the integral as a sum of infinitesimal quantities which represent the area below the graph of a curve, thus establishing the integral as an object in its own right. Of course, both Newton and Leibniz were describing essentially the same object, and the history of integration is, in some sense, the attempt to reconcile both definitions. A short overview of the early history of integration is given in Section 1.11 at the end of this chapter. For us, the modern theory of integration starts in the year 1854 with Riemann's Habilitationsschrift [143].

1.1 Modern Integration. 'Also zuerst: Was hat man unter $\int_a^b f(x)\,dx$ zu verstehen?'[1]

Riemann's answer to (t)his question is the following definition [143, Section 4]:

Definition 1.1 A bounded function $f : [a, b] \to \mathbb{R}$ defined on a compact interval $[a, b] \subseteq \mathbb{R}$ is **integrable** (in the sense of Riemann) if the limit

$$\int_a^b f(x)\,dx = \lim_{|\Pi|\to 0} \sum_{i=1}^{n} (x_i - x_{i-1}) f(\xi_i), \quad |\Pi| := \max_{1\leqslant i\leqslant n} |x_i - x_{i-1}|, \tag{1.1}$$

taken along all finite partitions $\Pi = \{a = x_0 < x_1 < \cdots < x_n = b\}$ and for any choice of intermediary points $\xi_i \in [x_{i-1}, x_i]$ exists and is finite.

[1] Riemann [143, p. 239] – To begin with: What is the meaning of $\int_a^b f(x)\,dx$?

Riemann immediately gives two necessary and sufficient conditions for the convergence of (1.1), cf. [MIMS, p. 443] for a modern proof. Denote by Π a finite partition of $[a, b]$ and write $D_i := \sup_{[x_{i-1},x_i]} f - \inf_{[x_{i-1},x_i]} f$ for the oscillation of a function f in the ith partition interval $[x_{i-1}, x_i]$.

(R1) The limit in (1.1) exists if, and only if, for all finite partitions Π of $[a, b]$,

$$\lim_{|\Pi|\to 0} \sum_{i=1}^{n} D_i \cdot |x_i - x_{i-1}| = 0.$$

(R2) The limit in (1.1) exists if, and only if,

$$\forall \epsilon > 0,\ \sigma > 0 \quad \exists \delta > 0 \quad \forall \Pi,\ |\Pi| \leqslant \delta \quad \sum_{i:D_i>\sigma} |x_{i-1} - x_i| < \epsilon.$$

In retrospect, Riemann's condition (R2) marks the beginning of the study of outer (Lebesgue) measure. We will see in Theorem 1.28 below that a bounded function f is Riemann integrable if, and only if, the set of its discontinuity points is a Lebesgue null set.

From (R2) it is clear that the Riemann integral is capable of dealing with functions which are discontinuous on a (countable) dense subset. This fact was already illustrated by Riemann in [143] using the function

$$f(x) = \sum_{n=1}^{\infty} \frac{h(nx)}{n^2}, \qquad h(x) = \begin{cases} x - k, & \text{if } x \in \left(k - \frac{1}{2}, k + \frac{1}{2}\right),\ k \in \mathbb{Z}, \\ 0, & \text{if } x = k \pm \frac{1}{2},\ k \in \mathbb{Z}, \end{cases}$$

which is discontinuous on the set $Q = \{p/(2n);\ \gcd(p, 2n) = 1\}$; see Fig. 3.3. Hankel [73, pp. 199–200] observed that f is an example of a function such that

$$F(x) := \int_0^x f(t)\,dt$$

is continuous, but $F'(x) = f(x)$ fails if $x \in Q$, i.e. F is not a primitive of f [☞ 3.7].

After the publication of Riemann's 1854 thesis in 1867, his definition of the integral was widely accepted, and it is still one of the most important and widely used notions of integration. The presentation was quickly streamlined, notably by the introduction of upper and lower sums and integrals which make Riemann's criterion (R1) more tractable.

Definition 1.2 (Thomae [183], Darboux [37], Volterra [195]) For a bounded function $f : [a, b] \to \mathbb{R}$ we call

$$S_\Pi[f] := \sum_{x_{i-1},x_i\in\Pi} m_i \cdot (x_i - x_{i-1}), \qquad m_i := \inf_{x\in[x_{i-1},x_i]} f(x),$$

$$S^{\Pi}[f] := \sum_{x_{i-1},x_i\in\Pi} M_i\cdot(x_i - x_{i-1}), \qquad M_i := \sup_{x\in[x_{i-1},x_i]} f(x),$$

the **lower** and **upper Darboux sums** and

$$\underline{\int_a^b} f(x)\,dx := \sup_{\Pi\subseteq[a,b]} S_{\Pi}[f] \quad \text{and} \quad \overline{\int_a^b} f(x)\,dx := \inf_{\Pi\subseteq[a,b]} S^{\Pi}[f]$$

(sup and inf range over all finite partitions Π of $[a,b]$) the **lower** and **upper Riemann–Darboux integrals**.

Using the lower and upper integrals we can show the following integrability criterion.

Theorem 1.3 ([MIMS, p. 443]) *A bounded function* $f : [a,b] \to \mathbb{R}$ *is Riemann integrable if, and only if,*

$$-\infty < \underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx < \infty.$$

The common (finite) value is the Riemann integral $\int_a^b f(x)\,dx$.

The development of the Riemann integral and the concept of a function go hand in hand. Up to Cauchy, functions were (implicitly) thought to be smooth, after Cauchy to be continuous; from 1867, Riemann integrable functions were seen to be the most general and still reasonable functions. But soon there were first examples of non-Riemann integrable functions, and other shortcomings of the Riemann integral were discovered:

1° The rather limited scope of Riemann integrable functions. The (proper) Riemann integral makes sense only on bounded sets and for bounded functions [☞3.2, 3.3], it behaves badly under compositions [☞3.11] and there are rather natural and simple non-integrable functions [☞3.4, 3.5].

2° If the Riemann integral is extended to two dimensions, the familiar formula

$$\iint_{[a,b]\times[c,d]} f(x,y)\,dx\,dy = \int_a^b \left[\int_c^d f(x,y)\,dy\right] dx = \int_c^d \left[\int_a^b f(x,y)\,dx\right] dy$$

may become senseless since some, or all, of the one-dimensional integrals might not exist [☞3.20].

3° Riemann's theory does not fix the difference between integral and primitive. There are integrable functions f such that $F(x) = \int_a^x f(t)\,dt$ is not everywhere differentiable, i.e. not a proper primitive. Worse, there are everywhere differentiable functions F whose derivative F' is not integrable [☞ 14.2, 14.3].

4° The Riemann integral behaves rather badly if one wants to interchange limits and integrals. Among other pathologies, one can construct a uniformly bounded sequence of Riemann integrable functions $(f_n)_{n\in\mathbb{N}}$ on $[0,1]$ such that $\lim_{n\to\infty} f_n(x) = f(x)$ but f is not Riemann integrable [☞ 3.13, 3.14, 3.16].

1.2 The Idea Behind Lebesgue Integration

Part of the problem with Riemann's definition is that the approximation procedure used in (1.1) is based on given partitions of the domain $[a,b]$ of the function $f : [a,b] \to \mathbb{R}$, i.e. these partitions need not relate to the behaviour of f.

Lebesgue's idea in [100, 101] is to split the range $f([a,b])$ of a bounded function $f : [a,b] \to \mathbb{R}$ into equal intervals, say $J_1, \dots, J_k$, and to determine those sets $I_1, \dots, I_k \subseteq [a,b]$ such that $I_i = f^{-1}(J_i)$. The corresponding approximations of the integral would be

$$U = \sum_{i=1}^{k} |I_i| \cdot \sup J_i \quad \text{and} \quad L = \sum_{i=1}^{k} |I_i| \cdot \inf J_i, \tag{1.2}$$

where $|A|$ denotes the total length of the set A. If we choose an equidistant partitioning of mesh δ, the value of the upper approximation is $U = L + \delta \cdot |[a,b]| = L + \delta \cdot (b-a)$, i.e. it is enough to restrict one's attention to L. Notice that the resulting partition of the domain depends on f. Before we give proper definitions and discuss the implications of this approach, let us consider a simple example.

Example 1.4 Consider an oscillating periodic function, e.g. $f(x) = \sin^2(n\pi x)$ with $n \in \mathbb{N}$ and $x \in [0,1)$, cf. Fig. 1.1. Using the relation $\sin^2\alpha = \frac{1}{2}(1 - \cos 2\alpha)$ it is easy to determine the integral of f,

$$\int_0^1 \sin^2(n\pi x)\,dx = \frac{1}{2}\int_0^1 (1 - \cos(2n\pi x))\,dx = \frac{1}{2}\int_0^1 dx = \frac{1}{2},$$

but the upper and lower Darboux sums for an equidistant partition of $[0,1]$, $\Pi = \{0 = x_0 < \cdots < x_k = 1\}$, with mesh $|\Pi| = \frac{1}{k} \geqslant \frac{1}{n}$ $\left(\text{or a general partition such that } \min_i(x_i - x_{i-1}) \geqslant \frac{1}{n}\right)$ are easily seen to be 1 and 0, respectively.

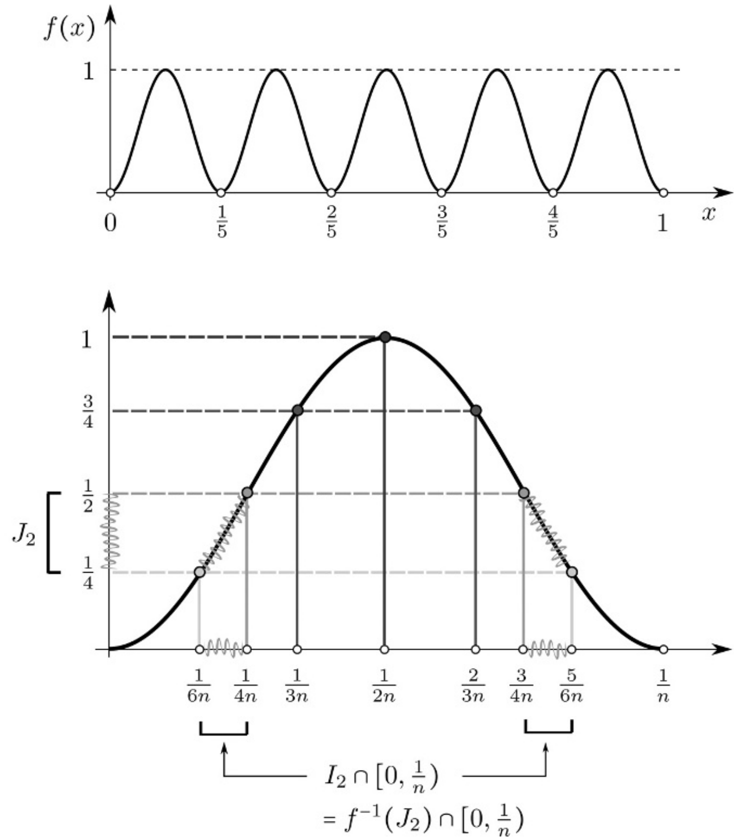


Figure 1.1 The oscillating periodic function $f(x) = \sin^2(n\pi x)$ for $n = 5$ (upper panel) and the choice of the domain partition I_k for an equidistant range partition J_k (lower panel).

However, Lebesgue's approach using $k = 4$ and $J_i = \left[\frac{i-1}{k}, \frac{i}{k}\right)$ gives, over the first period,

$$I_i \cap \left[0, \frac{1}{n}\right) = \left\{0 \leqslant x < \frac{1}{n}\,;\ \frac{i-1}{k} \leqslant f(x) < \frac{i}{k}\right\} = [x_{i-1}, x_i) \cup (x_{8-i}, x_{8-(i-1)}]$$

with $x_0 = 0$, $x_1 = 1/6n$, $x_2 = 1/4n$, $x_3 = 1/3n$, $x_4 = 1/2n$, $x_5 = 2/3n$, $x_6 = 3/4n$, $x_7 = 5/6n$ and $x_8 = 1/n$; see Fig. 1.1 (lower panel). Thus, in $[0, 1/n)$ we get

$$\sum_{i=1}^{4} \frac{i-1}{4}\left[(x_i - x_{i-1}) + (x_{8-(i-1)} - x_{8-i})\right] = \left(\frac{1}{24} + \frac{1}{12} + \frac{1}{4}\right)\frac{1}{n} = \frac{3}{8n}.$$

Since there are n periods in $[0, 1]$, we have $L = \frac{3}{8}$ (and $U = \frac{3}{8} + \frac{1}{4} = \frac{5}{8}$). This is already a reasonable approximation of the true value $\frac{1}{2}$ and, if one uses $\frac{1}{2}(U+L)$, it even happens to be the exact value.

This example makes it clear that Lebesgue's approach is better suited to deal with (rapidly) oscillating integrands, in particular, when the oscillations approach a condensation point as is the case for $x \longmapsto \sin^2 \frac{1}{x}$ as $x \to 0$.

1.3 Lebesgue Essentials – Measures and σ-Algebras

Let us recast Lebesgue's approximation of a function $f \geqslant 0$ from below by slicing its range **horizontally** as shown in Fig. 1.2. The level sets

$$A_k^n := \begin{cases} \{k2^{-n} \leqslant f < (k+1)2^{-n}\} & \text{for } k = 0, 1, 2, \dots, n2^n - 1, \\ \{f \geqslant n\} & \text{for } k = n2^n, \end{cases}$$

can be used to define step functions

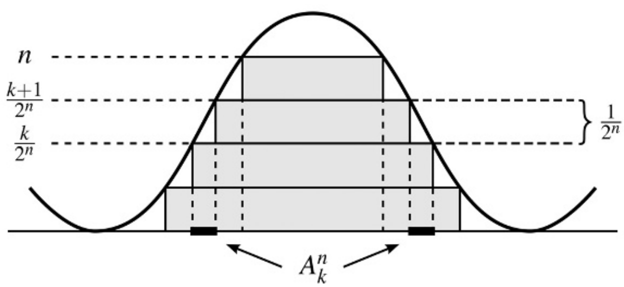


Figure 1.2 The function f sits like a 'Mexican hat' (a sombrero) over the approximating simple functions.

$$\phi_n(x) := \sum_{k=0}^{n2^n} k2^{-n} \mathbb{1}_{A_k^n}(x)$$

which approximate f from below. Coming from below has the advantage that f need not be bounded; instead, we use a moving cut-off level n which kicks in on the set $A_{n2^n}^n$. From Fig. 1.2 we see that

(i) $0 \leqslant \phi_n \leqslant \phi_{n+1} \leqslant f$ and $\phi_n \uparrow f$;
(ii) $|\phi_n(x) - f(x)| \leqslant 2^{-n}$ if $x \in \{f < n\}$; in particular, if f is bounded, the sequence ϕ_n approximates f uniformly.

We are interested in the nature of the level sets A_k^n. Property (i) requires that we are able to subdivide the level sets A_k^n finitely often. If we want to integrate the sum of two functions f, g, the level sets of $f + g$ will be expressed through finite unions and intersections of the level sets of f and g.

Therefore, the level sets form a family of sets which is closed if we repeat the usual set operations (intersection, union, taking complements and differences) finitely or – for limits – countably infinitely often. This requirement leads naturally to the notion of a σ-algebra.

Definition 1.5 Let $X \neq \emptyset$ be any set, and denote by $\mathscr{P}(X)$ its power set. A σ-**algebra** $\mathscr{A}$ on X is a family of subsets of X with the following properties:

$$X \in \mathscr{A}, \tag{Σ_1}$$

$$A \in \mathscr{A} \;\Rightarrow\; A^c \in \mathscr{A}, \tag{Σ_2}$$

$$(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{A} \;\Rightarrow\; \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}. \tag{Σ_3}$$

Because of (Σ_1) and (Σ_2) we have $\emptyset \in \mathscr{A}$, and using $A_1 \cup A_2 \cup \emptyset \cup \ldots$ in (Σ_3) shows that $\mathscr{A}$ is stable under finite unions. With de Morgan's laws we get that $\mathscr{A}$ is also stable under finite and countably infinite intersections and this is also true for differences as $A \setminus B = A \cap B^c$ is a combination of complementation and intersection.

The second ingredient needed for the construction of the integral is a 'gauge' for the size of the level sets A_k^n. In Example 1.4 we naively took the 'length' of the interval and there was no problem since the level sets were relatively simple. In the general case we need a function defined on all possible level sets which is compatible with (countably often repeated) set operations. This is the rationale for the following definition.

Definition 1.6 Let $X \neq \emptyset$ be any set. A (positive) **measure** is a set function $\mu\colon \mathscr{A} \to [0, \infty]$ satisfying

$$\mathscr{A} \text{ is a } \sigma\text{-algebra on } X, \tag{M_0}$$

$$\mu(\emptyset) = 0, \tag{M_1}$$

$$(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{A} \text{ pairwise disjoint} \;\Rightarrow\; \mu\left(\dot{\bigcup_{n\in\mathbb{N}}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n). \tag{M_2}$$

The pair $(X, \mathscr{A})$ is called a **measurable space** and $(X, \mathscr{A}, \mu)$ is called a **measure space**. The measure space is called **finite**, if $\mu(X) < \infty$, and σ-**finite**, if there exists a sequence $(F_n)_{n\in\mathbb{N}} \subseteq \mathscr{A}$ such that $X = \bigcup_{n\in\mathbb{N}} F_n$ and $\mu(F_n) < \infty$. A set $A \in \mathscr{A}$ is often called a **measurable set**.

The requirements (M_0)–(M_2) lead to a rich family of set functions with many further properties; see [MIMS, pp. 24, 28]. For example, if $A, B, A_n, B_n \in \mathscr{A}$:

(a) $A \cap B = \emptyset \Rightarrow \mu(A \cup B) = \mu(A) + \mu(B)$ (additive)

(b) $A \subseteq B \Rightarrow \mu(A) \leqslant \mu(B)$ (monotone)

(c) $A \subseteq B,\ \mu(A) < \infty \Rightarrow \mu(B \setminus A) = \mu(B) - \mu(A)$

(d) $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ (strongly additive)

(e) $\mu(A \cup B) \leqslant \mu(A) + \mu(B)$ (subadditive)

(f) $A_n \uparrow A \Rightarrow \mu(A) = \sup_n \mu(A_n) = \lim_n \mu(A_n)$ (continuous from below)

(g) $B_n \downarrow B,\ \mu(B_1) < \infty \Rightarrow \mu(B) = \inf_n \mu(B_n) = \lim_n \mu(B_n)$ (continuous from above)

(h) $\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) \leqslant \sum_{n\in\mathbb{N}} \mu(A_n)$. ($\sigma$-subadditive)

Example 1.7 Here are some of the most commonly used measures and σ-algebras. Unless otherwise indicated, $(X, \mathscr{A}, \mu)$ is an arbitrary measure space.

σ-Algebra $\mathscr{A}$	**Typical measure** on $(X, \mathscr{A})$
(a) The **indiscrete σ-algebra**: $\{\emptyset, X\}$ – this is the smallest possible σ-algebra on X.	▶ The **trivial measure** $\tau(\emptyset) = 0$ and $\tau(X) = \infty$.
(b) The **discrete σ-algebra**: $\mathscr{P}(X)$ – this is the largest possible σ-algebra on X.	▶ As a rule of thumb, rich σ-algebras admit only poor (i.e. simple) measures: $\mathscr{P}(X)$ can support the **trivial measure** from (a), the **counting measure** $\zeta(A) = \#A$, or **Dirac's delta function** (**point mass**) at $x \in X$ $\delta_x(A) = \begin{cases} 0, & \text{if } x \notin A, \\ 1, & \text{if } x \in A. \end{cases}$
(c) The **trivial σ-algebra**: $\mathscr{T}_\mu = \{A \in \mathscr{A}; \mu(A) = 0 \text{ or } \mu(A^c) = 0\}$.	▶ This construction works for every measure space $(X, \mathscr{A}, \mu)$.
(d) The **co-countable σ-algebra**: $\{A \subseteq X;\ \#A \leqslant \#\mathbb{N} \text{ or } \#A^c \leqslant \#\mathbb{N}\}$ on an uncountable set X [☞ Example 4A].	▶ The **co-countable** (probability) measure $\mu(A) = \begin{cases} 0, & \text{if } A \text{ is countable}, \\ 1, & \text{if } A^c \text{ is countable}, \end{cases}$ [☞ Example 5A].
(e) The **trace σ-algebra**: Let $E \subseteq X$. $\mathscr{A}_E = E \cap \mathscr{A} := \{E \cap A;\ A \in \mathscr{A}\}$.	▶ If $E \in \mathscr{A}$, the restriction $\mu\vert_E(A) := \mu(E \cap A)$ is a measure on the trace measurable space $(E, \mathscr{A}_E)$; [☞ 5.9] if $E \notin \mathscr{A}$.
(f) The **pre-image σ-algebra**: Let $\phi : X \to X'$ be any map and $\mathscr{A}'$ a σ-algebra on X'. $\phi^{-1}(\mathscr{A}') = \{\phi^{-1}(A');\ A' \in \mathscr{A}'\}$.	▶ If μ is a measure on $(X, \mathscr{A})$, then $\mu'(A') := \mu(\phi^{-1}(A'))$ is called the **image measure** or **push-forward measure** of μ under ϕ. Notation: $\mu\circ\phi^{-1}$, $\phi_*\mu$ or $\phi(\mu)$.

(g) The **σ-algebra generated by a set** $F \subseteq X$: This is the smallest σ-algebra on X containing the set F: $\sigma(F) = \{\emptyset, F, F^c, X\}$.

(h) The **σ-algebra generated by a family of sets** $\mathscr{F}$: This is the smallest σ-algebra containing the family $\mathscr{F}$: $\sigma(\mathscr{F}) = \bigcap\{\mathscr{B}\,;\ \mathscr{F} \subseteq \mathscr{B}, \mathscr{B}\ \sigma\text{-algebra}\}$.

(i) Let $\phi_i : X \to X_i$, $i \in I$ be arbitrarily many mappings and assume that $\mathscr{A}_i$ is a σ-algebra in X_i. The **σ-algebra generated by the family of mappings** $(\phi_i)_{i\in I}$, $\sigma(\phi_i, i \in I) = \sigma\left(\bigcup_{i\in I} \phi_i^{-1}(\mathscr{A}_i)\right)$, is the smallest σ-algebra that makes all ϕ_i measurable (see Definition 1.8 further on).

(j) The **completed σ-algebra**: Let $\mathscr{F} \subseteq \mathscr{A}$ be a (not necessarily proper) sub-σ-algebra,

$$\mathscr{N}_\mu = \{N \in \mathscr{A}\,;\ \mu(N) = 0\}$$

the family of all measurable null sets, and

$$\mathscr{N}_\mu^* = \{N^* \subseteq X\,;\ \exists N \in \mathscr{N}_\mu,\ N^* \subseteq N\}$$

the family of all subsets of measurable null sets.
The **completion** of $\mathscr{F}$ is the σ-algebra $\mathscr{F}^* := \sigma(\mathscr{F}, \mathscr{N}^*)$. One can show that

$$\begin{aligned}\mathscr{F}^* &= \{F \vartriangle N^*\,;\ F \in \mathscr{F},\ N^* \in \mathscr{N}^*\}\\ &= \big\{F^*\,;\ \exists A, B \in \mathscr{F},\ A \subseteq F^* \subseteq B,\\ &\qquad\qquad \mu(B \setminus A) = 0\big\}.\end{aligned}$$

▶ The **completion** $\bar{\mu}$ of the measure μ (defined on $\mathscr{F}$) is the measure $\bar{\mu}$ on the measurable space $(X, \mathscr{F}^*)$ given by

$$\bar{\mu}(F^*) := \frac{1}{2}(\mu(A) + \mu(B)), \quad F^* \in \mathscr{F}^*,$$

where the sets $A, B \in \mathscr{F}$ are such that $\mu(B \setminus A) = 0$ and $A \subseteq F^* \subseteq B$. The former ensures that $\bar{\mu}$ is well-defined, i.e. independent of the choice of the sets A and B.
Since $\mathscr{F} \subseteq \mathscr{F}^*$, $\bar{\mu}$ is an extension of μ.

(k) Let X be a topological space and $\mathscr{O}$ the family of all open sets. The **Borel** or **topological σ-algebra** is the σ-algebra generated by the open sets $\mathscr{B}(X) = \sigma(\mathscr{O})$.
Since a set is open if its complement is closed, $\mathscr{B}(X)$ is also generated by the closed sets. If X is a metric space which is the union of countably many compact sets $X = \bigcup_{n\in\mathbb{N}} K_n$ (e.g. if X is locally compact and separable), then $\mathscr{B}(X)$ is also generated by the compact sets [☞ 4.15, 4.16].

(l) The Borel sets in $\mathbb{R}^d$, $\mathscr{B}(\mathbb{R}^d)$, are generated by any of the following families: The open sets, the closed sets, the compact sets, the open balls $B_r(q)$ (radius $r \in \mathbb{Q}^+$, centre $q \in \mathbb{Q}^d$), the rectangles $\bigtimes_{i=1}^d [a_i, b_i)$ (with rational $a_i, b_i \in \mathbb{Q}$).

▶ Most measures used in analysis are defined on the Borel sets (or their completion, the Lebesgue sets, cf. Example (n)). The prime example of a measure on $\mathscr{B}(\mathbb{R}^d)$ is d-dimensional Lebesgue measure λ^d. Since the structure of the Borel sets is quite complicated, one defines λ^d on a sufficiently rich generator

$$\lambda^d\left(\bigtimes_{i=1}^d [a_i, b_i)\right) = \prod_{i=1}^d (b_i - a_i).$$

We will see in Theorem 1.40 that this characterizes λ^d uniquely.

(m) If $A \subseteq \mathbb{R}^d$, then $\mathscr{B}(A)$ is the Borel σ-algebra which is generated by the relatively open subsets of A. It is not hard to see that $\mathscr{B}(A)$ coincides with the trace σ-algebra $A \cap \mathscr{B}(\mathbb{R}^d)$.

▶ Use λ^d_A, the trace of Lebesgue measure λ^d on the trace-σ-algebra; see Example (e).

(n) The **Lebesgue σ-algebra** or **Lebesgue sets** $\mathscr{L}(\mathbb{R}^d)$ are the completion, see Example (j), of the Borel sets with respect to Lebesgue measure.

▶ Use the completion $\bar{\lambda}^d$ of λ^d; see Example (j).

(o) The **product σ-algebra** $\mathscr{A} \otimes \mathscr{B}$ is the σ-algebra $\sigma(\mathscr{A} \times \mathscr{B})$ generated by all generalized 'rectangles', i.e. sets of the form $A \times B \in \mathscr{A} \times \mathscr{B}$ with $A \in \mathscr{A}$ and $B \in \mathscr{B}$.

▶ Let $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ be σ-finite measure spaces. Similar to the construction of Lebesgue measure, the product measure ρ is defined first on the sets $A \times B \in \mathscr{A} \times \mathscr{B}$ of a generator,

$$\rho(A \times B) \coloneqq \mu(A)\nu(B),$$

and from the general theory it is known that this characterizes ρ on $\mathscr{A} \otimes \mathscr{B}$, cf. Theorem 1.33.

1.4 Lebesgue Essentials – Integrals and Measurable Functions

Let us return to the original problem of integrating a function. A real-valued function $f : X \to \mathbb{R}$ whose level sets $\{a \leqslant f < b\}$ are in a σ-algebra $\mathscr{A}$ on X is called measurable. The observation

$$\{a \leqslant f < b\} = \{f \geqslant a\} \cap \{f < b\} = f^{-1}([a, \infty)) \cap f^{-1}((-\infty, b))$$

explains the following slightly more general definition.

Definition 1.8 Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be two measurable spaces. A mapping $f : X \to Y$ is called $\mathscr{A}/\mathscr{B}$ **measurable**, if

$$\forall B \in \mathscr{B} \ : \ f^{-1}(B) \in \mathscr{A}. \tag{1.3}$$

If Y is a topological space equipped with its Borel sets, then measurable functions f are also called **Borel maps** or **Borel functions**.

Remark 1.9 (a) If $\mathscr{B}$ is generated by some family $\mathscr{H}$, then (1.3) is equivalent to the requirement that $f^{-1}(H) \in \mathscr{A}$ for all $H \in \mathscr{H}$. In particular, if we consider $\mathbb{R}$ equipped with the Borel σ-algebra $\mathscr{B}(\mathbb{R})$, then measurability of $f : X \to \mathbb{R}$ means that $\{f \leqslant a\} \in \mathscr{A}$ for all $a \in \mathbb{R}$ or $\{f > b\} \in \mathscr{A}$ for all $b \in \mathbb{R}$; see [MIMS, pp. 54, 60].

Since the pre-image of an open set under a continuous function is open, continuous functions are always Borel measurable.

(b) The pre-image σ-algebras from Example 1.7.(f), (i) are the smallest σ-algebras in X such that the map $\phi : X \to X'$, resp. all maps $\phi_i : X \to X_i$, become measurable [MIMS, p. 55].

(c) Let $(X, \mathscr{A})$, $(Y, \mathscr{B})$ and $(Z, \mathscr{C})$ be measurable spaces and $f : X \to Y$ and $g : Y \to Z$ be $\mathscr{A}/\mathscr{B}$, resp. $\mathscr{B}/\mathscr{C}$, measurable maps. The composition $f \circ g$ is $\mathscr{A}/\mathscr{C}$ measurable [MIMS, p. 54].

(d) Let $(Z, \mathscr{C})$ and $(X_i, \mathscr{A}_i)$, $i = 1, 2$, be measurable spaces. The product σ-algebra, see Example 1.7.(o), is constructed in such a way that a mapping $f = (f_1, f_2) : Z \to X_1 \times X_2$ is $\mathscr{C}/\mathscr{A}_1 \otimes \mathscr{A}_2$ measurable if, and only if, the coordinate maps $f_i : Z \to X_i$ are $\mathscr{C}/\mathscr{A}_i$ measurable [MIMS, p. 149]. This construction is analogous to the definition of an initial topology (also known as the weak or limit or projective topology).

(e) The family of Borel measurable real-valued functions is a vector space which is closed under countable pointwise infima and suprema. In particular, the lim inf, lim sup and lim of a sequence of Borel measurable functions is again a Borel measurable function – possibly with values in the extended real line $[-\infty, \infty]$; see [MIMS, pp. 66, 67]. ☞ 8.15-8.18

A **simple function** is a function $\phi : X \to \mathbb{R}$ of the form

$$\phi(x) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}(x), \quad \alpha_i \in \mathbb{R},\ A_i \in \mathscr{A}. \tag{1.4}$$

Without loss we can assume that $A_i = \{\phi = \alpha_i\} = \{x \in X\,;\ \phi(x) = \alpha_i\}$. We write $\mathcal{E} = \mathcal{E}(X)$ for the family of all simple functions on a measurable space $(X, \mathscr{A})$. A simple function is obviously measurable and attains at most finitely many values. The following characterization of measurability of real functions is important for the development of the integral. It is, at the same time, an important structural result for the class of measurable functions. Its proof is, essentially, summed up in Fig. 1.2.

Theorem 1.10 (sombrero lemma [MIMS, pp. 64, 65]) *Let $(X, \mathscr{A})$ be any measurable space. A positive function $f : X \to [0, \infty]$ is $\mathscr{A}/\mathscr{B}(\overline{\mathbb{R}})$ measurable if, and only if, there exists an increasing sequence of simple functions $\phi_n \geqslant 0$ such that $f = \sup_{n \in \mathbb{N}} \phi_n$. If f is bounded, then ϕ_n approximates f uniformly.* ☞ 8.22

Remark 1.11 In measure theory one often works with the **extended real line** $\overline{\mathbb{R}} = [-\infty, \infty]$ which is the two-point compactification of $\mathbb{R}$ [☞ Example 2.1(f)]; in addition to the usual neighbourhoods of points $x \in \mathbb{R}$, we have the neighbourhoods of $\pm\infty$ which are all sets containing $(x, +\infty]$, resp. $[-\infty, y)$, for some $x, y \in \mathbb{R}$. The Borel sets are $\mathscr{B}(\overline{\mathbb{R}}) = \sigma(\mathscr{B}(\mathbb{R}), \{\infty\}, \{-\infty\})$ and $\mathscr{B}(\mathbb{R})$ is the trace $\mathbb{R} \cap \mathscr{B}(\overline{\mathbb{R}})$.

It is clear how to add or multiply $x, y \in [-\infty, +\infty]$ unless both $x, y = \pm\infty$. For these cases, we agree that

$$\infty + \infty := \infty, \quad -\infty - \infty := -\infty, \quad \text{and} \quad 0 \cdot (\pm\infty) := 0,$$

while expressions of the type '$\infty - \infty$', '$-\infty + \infty$' or '$\pm\infty/\infty$' **are not defined**. Unless otherwise stated, we equip $\mathbb{R}$ and $\overline{\mathbb{R}}$ with the Borel σ-algebra $\mathscr{B}(\mathbb{R})$ and $\mathscr{B}(\overline{\mathbb{R}})$.

If we combine the idea that the integral represents the area below the graph of a positive function with the sombrero lemma, we naturally arrive at the following definition.

Definition 1.12 Let $(X, \mathscr{A}, \mu)$ be a measure space.

(a) The **integral** of a real-valued simple function $\phi \in \mathcal{E}(X)$ is

$$I_\mu(\phi) := \sum_{\alpha \in \phi(X)} \alpha \mu\{\phi = \alpha\} \in \mathbb{R}. \tag{1.5}$$

(b) The **integral** of a positive measurable function $f : X \to [0, \infty]$ is

$$\int f \, d\mu := \int f(x) \, \mu(dx) := \sup\left\{I_\mu(\phi) ; \ \phi \in \mathcal{E}(X), \ 0 \leqslant \phi \leqslant f\right\} \in [0, \infty]. \tag{1.6}$$

It is obvious that both (1.5) and (1.6) are well-defined, i.e. independent of the representation, resp. approximation. Since the (finite) sum in (1.5) is linear and monotone, this is also true for the functional $\phi \longmapsto I_\mu(\phi)$. While it is easy to see that (1.6) extends (1.5) – i.e. $\int \phi \, d\mu = I_\mu(\phi)$ – it is not clear that the supremum preserves linearity. The following theorem comes to the rescue.

Theorem 1.13 (Beppo Levi [MIMS, p. 75]) *Let $(X, \mathscr{A}, \mu)$ be a measure space and $(f_n)_{n\in\mathbb{N}}$ an increasing sequence $0 \leqslant f_n \leqslant f_{n+1} \leqslant \dots$ of positive measurable functions. The limit $f := \sup_{n\in\mathbb{N}} f_n : X \to [0, \infty]$ is measurable and*

$$\int \sup_{n\in\mathbb{N}} f_n \, d\mu = \sup_{n\in\mathbb{N}} \int f_n \, d\mu. \tag{1.7}$$

If $f : X \to [0, \infty]$ is measurable, then the sombrero lemma (Theorem 1.10) tells us that there is an increasing sequence of positive simple functions such that $\phi_n \uparrow f$. Therefore, the supremum $\int f \, d\mu$ is, in fact, an increasing limit $\lim_{n\to\infty} I_\mu(\phi_n)$. This means that on the positive measurable functions the functional $f \longmapsto \int f \, d\mu$ is monotone, additive and positively homogeneous.